

[illegible]

XX WT1: 2802 984314712.
DR P-PSDB: AAB47837.
XX
XX New isolated chimeric receptor comprising a DNA binding domain and/or
PT 11and binding domain of a new estrogen receptor, for identifying
XX functional ligands or hormonal analogs for the receptor.
XX
XX Example B; Page 13; 35pp; English.

The sequences given in AAI121146 and AAI121159 encode splice variants
CC of a novel estrogen receptor (ER). The gene encoding this new ER is
CC located on chromosome 14 and has a different tissue distribution from
CC classical ER. This ER also has two orphan ER's, ER alpha and ER beta.
CC These orphan receptors have estrogen receptor related structure but do
CC not appear to be able to bind estradiol or other ER ligands. These
CC splice variants differ from the parent ER around exon 8. One clone
CC contains exon 8b through alternative splicing. This causes an immediate
CC termination of the reading thereby creating a truncation at the carboxy
CC terminus. The other splice variant contains an alternative exon 8, exon
CC 8c, which encodes two C-terminal amino acids and then has a stop codon.
CC The splice variant proteins do not contain the AF-2 region and therefore
CC probably lack the ability to modulate transcription of target genes in a
CC ligand dependent manner. However they may be able to interfere with
CC the functioning of the WT classical ER or the novel ER of the invention,
CC either by heterodimerization or by occupying estrogen response elements
CC or by interactions with other transcription factors.

XX
XX Sequence 1251 BP; 297 A; 324 C; 354 G; 276 T; 0 other;

Query Match	100.0%	Score 1251;	DB 24;	Length 1251;
Match 1251;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY 1 ATGATATACAGCATTCCTCCAGCAATCTCACTAACTGTGAAGTGTGATCTGTGATCAAGC 60				
DB 1 ATGATATATACGATTCCTCCAGCAATTCATCACTGTGAAGTGTGATCTGTGATCAAGC 60				
QY 61 ACAAGGCCAATGTTGTTGGCCACACCTGAGGACCTTTCTCTTTAGTGGTCAATGC 120				
DB 61 ACAAGGCCAATGTTGTTGGTGGACACCTGAGGACCTTTCTCTTTAGTGGTCAATGC 120				
QY 121 CAGTTATCAGCATCTGTAATCGGACACCTCAAAAGATGCTCTGTGTGAAACAAGATGCTA 180				
DB 121 CAGTTATCAGCATCTGTAATCGGACACCTCAAAAGATGCTCTGTGTGAAACAAGATGCTA 180				
QY 181 GAGCAACCTTACTTCTTAAACATGACAGCAATCTGAAACAAACATTAATCTGCAAACTTCT 240				
DB 181 GAGCAACCTTACTTCTTAAACATGACAGCAATCTGAAACAAACATTAATCTGCAAACTTCT 240				
QY 241 GCGAGGCTGTGTATCTGTCCAGTTCGAAGAGAGGAGATCTTACTCTGTGCTGTGAGG 300				
DB 241 GCGAGGCTGTGTATCTGTCCAGTTCGAAGAGAGGAGATCTTACTCTGTGCTGTGAGG 300				
QY 301 GATTACGATCGGATATATCATCATATGAGATGTGGTGTGAAGGATCTTAAGGACTTTT 360				
DB 301 GATTACGATCGGATATATCATCATATGAGATGTGGTGTGAAGGATCTTAAGGACTTTT 360				
QY 361 AAAAGAAACATTTAAAGACATAATGATTATATTGTCTGAGCTACAAATGAGTGAATC 420				
DB 361 AAAAGAAACATTTAAAGACATAATGATTATATTGTCTGAGCTACAAATGAGTGAATC 420				
QY 421 GATTAAGGCGGAGGAGCAAGCTGTGAGGCTGTGGCATCTGGAAAGTGTATGAAATGTGGGA 480				
DB 421 GATTAAGGCGGAGGAGCAAGCTGTGAGGCTGTGGCATCTGGAAAGTGTATGAAATGTGGGA 480				
QY 481 ATGTGTAAGTGTATCTGTTTATATAGAGAGATGTGGTATATGCTGTGTGTGCAATCAACA 540				
DB 481 ATGTGTAAGTGTATCTGTTTATATAGAGAGATGTGGTATATGCTGTGTGTGCAATCAACA 540				
QY 541 AGTGTGTAGCAGAGAGCTGTGATCTGTGTGGGTGAAGGTCAAGATCAATGTGTGACATG 600				
DB 541 AGTGTGTAGCAGAGAGCTGTGATCTGTGTGGGTGAAGGTCAAGATCAATGTGTGACATG 600				

UY	"4	AGTCGCAACAACTATGTACTGGTCCGGAAGAAGCATTAAGAGAAATGGGGCGGCATCCTCC	600
DB	1118	ATTGCGTCAAGAGATGTTGATCTGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1177
UY	601	CCTACCTGGAGAGCTGGCTGGAG	660
DB	1178	GCAATGTGGAGAGATGCTGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1237
UY	661	GAGCTGAG	720
DB	1238	GAGCTGAG	1297
UY	721	ATGAT	780
DB	1298	ATGAT	1357
UY	781	AAGCAAT	840
DB	1358	AAGCAAT	1417
UY	841	TCTTGAGAT	900
DB	1418	TCTTGAGAT	1477
UY	901	CTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	960
DB	1478	CTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1537
UY	961	CTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAAT	1020
DB	1538	CTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAATCTTCAAAAT	1597
UY	1021	CATCAAAAT	1080
DB	1598	CATCAAAAT	1657
UY	1081	GTCAG	1140
DB	1658	GTCAG	1717
UY	1141	ATCATATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1200
DB	1718	ATCATATCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1777
UY	1201	GCTGTGGCTAAT	1247
DB	1778	GCTGTGGCTAAT	1824
RESULT 15			
AHNH6L25 standard: 100%; 244 bp.			
AHNH6L25:			
XX	23 SEP 2002 (first entry)		
DE	Nucleotide sequence of a fusion GFP/bkx beta gene.		
XX			
RW	Fluorescent analysis: laser? fluorescence; GFP; FR beta; human? gene?		
XX	green fluorescent protein; estrogen receptor beta; fusion; ds.		
OS	Synthetic.		
OS	Acquired Victoria.		
XX	Bmo sapiens.		
XX			
FM	Key	local ion/unalitions	
FT	CD5	B...2389	
FT	/seq	a	
XX	/product	"GFP/bkx beta fusion protein"	
XX	Wc200249694 AL		
XX			

19	20 JUN-2002.
XX	
XX	14 DEC-2001; 2001MO-JP10998.
XX	
XX	14-DEC-2000; 2000JP-0380427.
XX	30-JAN-2001; 2001JP-0322105.
XX	
EA	(GLYH) OLYMPUS OPTICAL CO LTD.
XX	
P1	Kato N, Sakamoto H;
XX	
D8	WP1: 2002-528176/56.
D8	P-PSDB; ABH0756.
XX	
P1	Autocorrection function embedded confocal optics based fluorometric
P1	analyzers for studying behaviors of fluorescence-labeled molecules
P1	particularly intracellular biological molecules like proteins at
P1	molecular level .
XX	
PS	Example 2; Page 98-105; 10pp; Japanese.
XX	
CC	The invention relates to a fluorometric analyzer that comprises a laser
CC	light source, an optical system for focusing the light beam from laser
CC	light source to a sample to form a confocal region, another optical
CC	system for focusing fluorescence from the sample, a light detector for
CC	measuring intensity, and a recording means to record the variation of
CC	intensity in the measured fluorescence. The analyzers are for studying
CC	behaviors of fluorescence-labeled molecules particularly intracellular
CC	biological molecules like proteins at molecular level, e.g. protein
CC	functions and interactions. Such analyzers are stable and convenient to
CC	operate, thereby enabling easy performance of fluorescence correlation
CC	spectroscopy, fluorescence intensity distribution analysis and
CC	fluorescence intensity multiple distribution analysis. The present
CC	sequence represents the nucleotide sequence of a fusion gene comprising
CC	a GFP (green fluorescent protein) gene and hER (human estrogen receptor)
CC	beta gene.
XX	
SU	Sequence 2433 BP; 587 A; 704 C; 659 G; 483 T; 0 other;
Query Match:	99.7%; Score 1247; IB 24; Length 2433;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1247:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGAAATTAAAGATTCGCCAGCAATGTACTAACTTGGAAGTGGGCTGGTGACAACT 60
DB	956 ATGAATTAAAGATTCGCCAGCAATGTACTAACTTGGAAGTGGGCTGGTGACAACT 1014
QY	61 ATTAAGCGCAAAATGCTGTGGGCACAACACTGACACTCTTCTTACTGGCAGCAACT 120
DB	1016 ATTAAGCGCAAAATGCTGTGGGCACAACACTGACACTCTTCTTACTGGCAGCAACT 1073
QY	121 CAATTATACAACTTCCTATGCGCGAATCTCAAAAAGTCTCTGGTGCAAGTAAGATG 180
DB	1076 CAATTATACAACTTCCTATGCGCGAATCTCAAAAAGTCTCTGGTGCAAGTAAGATG 1144
QY	181 GAAGAATCTTTACGTGAAGACAGAACAACTATAAAGCAAGCTTAGTGGCAAGCTGGC 240
DB	1146 GAAGAATCTTTACGTGAAGACAGAACAACTATAAAGCAAGCTTAGTGGCAAGCTGGC 1193
QY	241 GTCTAGCTCTGTTATCTGGCTCAAGTTCAAAACAGCATATCTATCTGGATCTGGTAC 300
DB	1196 GTCTAGCTCTGTTATCTGGCTCAAGTTCAAAACAGCATATCTATCTGGATCTGGTAC 1255
QY	301 GATTACGATCGGATATCACTATGGAATCTGGCTGTGGAAAGATGAAAGCTTTTT 460
DB	1256 GATTACGATCGGATATCACTATGGAATCTGGCTGTGGAAAGATGAAAGCTTTTT 1415
QY	361 AAAAAGAAATATTCAAAGACATAATGATATATTTGGTCAAGTAAATAGATGAATATC 420
DB	1416 AAAAAGAAATATTCAAAGACATAATGATATATTTGGTCAAGTAAATAGATGAATATC 1479
QY	421 GATAAAAAATCTGGCTGAAGAGCTGGCAAGCTTGGCTGCAAGTGTATCGAAGTGAATA 480

QY 1075 CTCTGGTACAGAGTACATGATGCT---GACAGAGCGGAGATGAGTACTGCTG 1141
 DB 1843 ACATTTCTGTCCAGTACGTCAGAGTCTGTGGAGAGAGACACATATCAACAGAGTCTG 1902
 QY 1132 AACGGGATGACGATGCTTTGCTTGGGATGATGACCAAGCGACATCTCTGGAGCAG 1191
 DB 1903 GACAAATGACAGACACTTTATGTCAGCTGATGATGACAGAGGCTGACCTTGAGTACAG 1962
 QY 1192 CAATTCATGCTGCTTACTTAACTCTGATGCTGCTGCTGACATGACCATGCGAG 1247
 DB 1963 CAGGACCAATGAGCTGGGCTGAGTCTGCTGATGCTGCTGACATGAGGACATGAG 2018
 RESULT 38
 AA86920
 ID AA86920 standard: cDNA: 2092 bp.
 AC AAC86920:
 DT 02-APR-2001 (first entry)
 XX
 DB Nucleotide sequence of the human estrogen receptor cDNA
 XX
 KM Ribozyme; estrogen dependent tumour; cell proliferation; glucocorticoid;
 KM DNA-binding domain; estrogen receptor; cancer treatment; breast cancer;
 KM ss.
 XX Homo sapiens.
 OS
 FN W0200074485-A1.
 XX
 PD 14-DEC-2000.
 XX
 PE 02-JUN-2000: 2000W0-US15243.
 XX
 PR 04-JUN-1999: 9908-0137470.
 XX
 PA (TEXA) UNIV TEXAS.
 XX
 PI Roy AK, Lavetskyy Y, Tyagi RK, Song GS, Chatterjee B.
 XX
 DB W01: 2001 061633/97.
 XX
 PT Ribozyme having a high substrate specificity for an mRNA encoding a
 PT DNA binding domain of human estrogen receptor, useful for inhibiting a
 PT estrogen dependent tumour cell proliferation, particularly breast cancer
 PI
 XX Disclosure: Page 8-9; 4pp; English.
 XX
 PS The specification describes a ribozyme capable of inhibiting
 CC estrogen-dependent tumour cell proliferation and having a high
 CC substrate specificity for an mRNA sequence encoding a DNA-binding
 CC domain of human estrogen receptor. The ribozyme is free of endonuclease
 CC activity for an mRNA having a DNA binding domain of a glucocorticoid.
 CC The estrogen receptor site-specific ribozymes are useful for cancer
 CC treatment and therapies, especially for inhibiting estrogen-dependent
 CC tumour cell proliferation, particularly breast cancer. The present
 CC sequence represents the human estrogen receptor cDNA.
 XX
 SU Sequence 2092 BP: 473 A; 605 C; 593 G; 421 T; 0 other:
 Query Match 25.6%, Score 320.4, E-Val 2e-22, Length 2072;
 Host Local Similarity 60.6%; Pred. No. 3,20-74;
 Matches 616; Conservative 0; Mismatches 361; Indels 39; Gaps 4;
 QY 271 AGGAGATCTTACTTCTGCTGCTGCTGAGGATTAAGATGAGATGATGAGATC 330
 DB 830 AAGGAGATCTTACTTCTGCTGCTGAGGATTAAGATGAGATGATGAGATC 889
 QY 331 TGGTGGTGAAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 490
 DB 1111 1111 11 11 111111 11 1111 111111111111 11 111

DB 890 TGGTGGTGAAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 949
 QY 491 ATTGTGACAGTACCAATAGTGTATATATATATATATATATATATATATATATAT 450
 DB 950 ATGTGTGACAGTACCAATAGTGTATATATATATATATATATATATATATATAT 1009
 QY 451 TTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 510
 DB 1010 TTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1069
 QY 511 TTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 558
 DB 1070 AATGAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 1129
 QY 559 CACTGTGCGGAG 607
 DB 1130 GTGGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1184
 QY 608 -----GGGAGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
 DB 1190 GTGGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248
 QY 658 GTGGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
 DB 1250 TTGAGTGTGAG 1309
 QY 715 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
 DB 1410 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
 QY 775 TGGGCTGAG 834
 DB 1370 TGGGCTGAG 884
 QY 835 GAGAGCTGTGATGAG 894
 DB 1430 GAGAGCTGTGATGAG 1489
 QY 895 GAGAGCTGTGATGAG 954
 DB 1440 GAGAGCTGTGATGAG 1499
 QY 955 GAGAGCTGTGATGAG 1014
 DB 1550 GAGAGCTGTGATGAG 1069
 QY 1015 GAGAGCTGTGATGAG 1074
 DB 1610 GAGAGCTGTGATGAG 1669
 QY 1075 GAGAGCTGTGATGAG 1131
 DB 1670 GAGAGCTGTGATGAG 1726
 QY 1132 AACGGGATGACGATGCTTTGCTTGGGATGATGACCAAGCGACATCTCTGGAGCAG 1191
 DB 1730 GAGAGATGACGATGCTTTGCTTGGGATGATGACCAAGCGACATCTCTGGAGCAG 1789
 QY 1192 CAATTCATGCTGCTTACTTAACTCTGATGCTGCTGACATGACCATGCGAG 1247
 DB 1790 CAATTCATGCTGCTTACTTAACTCTGATGCTGCTGACATGACCATGCGAG 1845
 RESULT 39
 AB151900
 ID AB151900 standard: cDNA: 2092 bp.
 AC AB151900:
 DT 10-JUL-2002 (first entry)
 XX
 DB Human estrogen receptor encoding cDNA SPO ID NO:9.
 XX
 DE



UY 983 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1042
1111 11 11 1111 11 11 11 11 11 11 11 11 11 11 11 11
DB 1020 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1879
UY 1043 AGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1101
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1080 AATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1949
UY 1102 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1159
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1140 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1999
UY 1160 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1219
1111 1111 11 11 11 1111 11 11 1111 11 11 1111 1111
DB 2000 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 2059
UY 1220 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 1247
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2060 TGGTAATCTTAAAGTTTGGAGCTTAAAGCTTCAACAAAGATACTGCTGCA 2087

RESULT 16

US-09-040-508-1

Sequence 1, Application US/09040508

Patent No. 6069289

GENERAL INFORMATION:

APPLICANT: Mathias, Stephen L.

TITLE OF INVENTION: NOVEL COMPOUNDS

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: KATNER & PRESTIA

STREET: P.O. BOX 980

CITY: VALLEY FORGE

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/040, 508

FILING DATE: 17 MAR 1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9705451.4

FILING DATE: 17 MAR 1997

APPLICATION NUMBER: GB 9803289.9

FILING DATE: 16 FEB 1998

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F

REGISTRATION NUMBER: 24,041

REFERENCE/BOOKET NUMBER: GB-40008

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610 407-0780

TELEFAX: 610 407-0701

TELEX: 845169

INFORMATION FOR SEQ ID NOS: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1615 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLGY: linear

MOLECULE TYPE: cDNA

US-09-040-508-1

Query Match 9, 18; Score 113.8; DB 3; Length 1615;

Post Local Similarity 49.38; Prod. No. B.26-215;

Matches 481; Consensitive 0; Mismatches 562; Indels 30; Gaps 2;

UY 292 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 451

DB 545 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 604
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
UY 452 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 411
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 605 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 664
UY 412 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 471
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 665 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 724
UY 472 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 531
111111 11 1111 11 11 11 11 11 11 11 11 11 11 11 11
DB 725 AAGGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 763
UY 532 AAGGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 591
764 -----GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 817
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 592 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 651
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 818 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 877
UY 652 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 711
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 878 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 937
UY 712 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 771
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 938 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 997
UY 772 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 831
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 998 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1057
UY 832 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 891
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1058 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1117
UY 892 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 951
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1118 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1174
UY 952 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1011
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1175 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1244
UY 1012 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1064
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1245 GTGTGAGGATTAAGATGAGATATGAGTCTCTCTGTCGAGAGATATGAG 1287

RESULT 17

US-09-500-654-1

Sequence 1, Application US/09500654

Patent No. 6359116

GENERAL INFORMATION:

APPLICANT: Stephen L. Mathias

TITLE OF INVENTION: No. 6359116el compounds

FILE REFERENCE: GP-30008-D1

CURRENT APPLICATION NUMBER: US/09/500, 654

CURRENT FILING DATE: 2000-02-09

EARLIER APPLICATION NUMBER: 9705451.4

EARLIER FILING DATE: 1997-03-17

EARLIER APPLICATION NUMBER: 9803289.9

EARLIER FILING DATE: 1998-02-16

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 1615

TYPE: DNA

ORGANISM: HOMO SAPIENS

US-09-500-654-1


```

1 CLASSIFICATION: 435
2 ALTERNATE/AGENT INFORMATION:
3 NAME: KILL, Gordon
4 REGISTRATION NUMBER: 30,764
5 REFERENCE/JACKET NUMBER: A-6612
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: (292) 293-7060
8 TELEFAX: (292) 293-7860
9 INFORMATION FOR SEQ ID NO: 1:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 3014 base pairs
12 TYPE: nucleic acid
13 STRANDEDNESS: single
14 TOPOLOGY: linear
15 MOLECULE TYPE: cDNA
16 HYDROTHERICAL: NO
17 US-08-629-939-1
18
19 Query Match 6.5% Score 81; DB 1; Length 3014;
20 Best Local Similarity 63.7%; Pred. No. 5,36-12;
21 Matches 123; Gaps 0; Mismatches 79; Indels 0; Gaps 0;
22
23 QY 293 TCTTCACGATTAAGTATCGCATATCACTATGCATTGTCTTTAAACTATAGC 352
24 ||||| - - - | - - - || - - - || - - - || - - - || - - - || - - - ||
25 1881 TCCTGGTGGAATCAAGCATATGAGCTGTGATTAATGATGCTTACCTGTAAGG 1940
26
27 CY 353 CTTTTTAAAAGAACATCAAGGACATAAATATATTTTTCAGCTACAAGAAT 412
28 ||||||| - - - | - - - || - - - || - - - || - - - || - - - || - - - ||
29 1941 TCTTCTTAAAGGGAATGAAAGGACAGCAACAACTACTATATGCTGTAAGAAAGCT 2000
30
31 CB 413 CATAAATCAAAAAGGAGGAGTAAAGCTCTCAAGGCTGTCACCTTGGAACTTACG 472
32 ||||| - - - | - - - || - - - || - - - || - - - || - - - || - - - ||
33 2001 CATATCTTCAAAAAAGGAGGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2060
34
35 CB 473 AACGCGAAATGCT 485
36 || - - - || - - - || - - - || - - - || - - - || - - - || - - - ||
37 2061 AACGCGAAATGCT 2073
38
39 RESULT 27
40 US-08-759-873-1
41 Sequence 1, Application US/08759873
42 Patient No. 5681885
43 GENERAL INFORMATION:
44 APPLICANT: Kieback, Dirk G.
45 TITLE OF INVENTION: METHODS FOR DIAGNOSING AN INCREASED RISK
46 TITLE OF INVENTION: OF BREAST OR OVARIAN CANCER
47 NUMBER OF SEQUENCES: 14
48 CORRESPONDENCE ADDRESS:
49 ADDRESSEE: SUGRONE, MICHAEL ZINN, MACPHEAR & SEARS
50 STREET: 2100 Pennsylvania Avenue, N.W., Suite 800
51 CITY: Washington, D.C.
52 STATE: D.C.
53 COUNTRY: U.S.A.
54 ZIP: 20037
55 COMPUTER READABLE FORM:
56 MEDIUM TYPE: Floppy disk
57 COMPUTER: IBM PC compatible
58 OPERATING SYSTEM: PC-DOS/MS-DOS
59 SOFTWARE: Patent In Release #1.0, Version #1.25
60 CURRENT APPLICATION DATA:
61 APPLICATION NUMBER: US/08/759,873
62 FILING DATE: 12-APRIL-1996
63 CLASSIFICATION: 435
64 ATTORNEY/AGENT INFORMATION:
65 NAME: KILL, Gordon
66 REGISTRATION NUMBER: 30,764
67 REFERENCE/JACKET NUMBER: A-6612
68 TELECOMMUNICATION INFORMATION:
69 TELEPHONE: (292) 293-7060
70 TELEFAX: (292) 293-7860
71 INFORMATION FOR SEQ ID NO: 1:
72 SEQUENCE CHARACTERISTICS:

```

QY	284	TTGGGCTGTCTGCAAGCAATTACGATGGAATATCAATATGAGATGTGGTGTGTGGAG	443
Db	785	TGTGGCTGTCTGTGCAAGCAATTACGATGGAATATCAATATGAGATGTGGTGTGTGGAG	844
QY	444	GATGTAAGGCGCTTTTAAAAACAGCAATTTAAACAGCAATATGATTTGCTGCACTA	403
Db	845	GCTGTCAAGGCTTTCTTAAAGTGCATGCTGTCAAAAAAAGTCAAAATATGCTGTGTAA	904
QY	404	CAAAATAGTGTAAATCAATCAAAAAACGAGGCGAAGATTCGTAAGGCTGTCACTTCGAA	463
Db	905	ATAAAAATTTGCGTACAGTGAATAAGTGGCGGCGCAAAATGCTTTTCATATGCTGCTTCA	964
QY	464	AGTGTTCACAGATTCGAAATGCTTGAGGTGTGCTGTGGAATGAAGATAT	511
Db	965	AGTGTTCGATTCGTGGAGATGCTTAAACAAAGTGGTGTGGAATGAAGATAT	1012

RESULT 31
US-08-592-483-4
Sequence 3, Application US/09572383
Patent No. 5840760
GENERAL INFORMATION:
APPLICANT: Isai, S. and S.J. Collins
TITLE OF INVENTION: "Hematopoietic Cell Lines Bearing Altered Retinoid Acid Recept
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STREET: 2800 Pacific First Centre, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-5.25 inch, 1.2MB storage
OPERATING SYSTEM: MS-DOS 4.01
SOFTWARE: Word for Windows 5.01-1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,383
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/099,242
FILING DATE: July 28, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,432
REFERENCE/DOCKET NUMBER: FHP-1-7190
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-622-8100; 1-206-224-0700 (direct)
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2658 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: page 4, KAK-alpha403 dominant negative; deleted of 1411-1504 of SEQ.
IMMEDIATE SOURCE:
LIBRARY: cDNA
US-08-592-483-4
Query Match 5.9%, Score 74.4; DB 2; Length 2658;
Best Local Similarity 60.3%; Pred. No. 2, 80-10;
Matches 123; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
DB 286 TGGCTGTGACGATTAAGCATGAGATATGAGAGTGTGTGTGAGAGA 345
DB 364 TGGTTGTGTAAGAAATGCTGAACTAGTATGAGAGTGTGTGTGAGAGA 424
DB 346 TGAAGGCTTTTAAAGAGCATTAAGGACATATGATATTTGCGAGTACA 405
DB 424 TGAAGGCTTTTAAAGAGCATTAAGGACATATGATATTTGCGAGTACA 483
DB 406 AATCAGTGTACATATTAATAAAGAGGAGTGTGTGTGAGAGTGTGTGAG 465
DB 484 AATCAGTGTACATATTAATAAAGAGGAGTGTGTGTGAGAGTGTGTGAG 544
DB 466 TGTTCGAGAGTGTGTGTGAGAGTGTGTGTGAGAGTGTGTGTGAGAG 489
DB 544 TGTTCGAGAGTGTGTGTGAGAGTGTGTGTGAGAGTGTGTGTGAGAG 567

RESULT 32
US-08-095-728B-3
Sequence 3, Application US/0806728B
Patent No. 5841642
GENERAL INFORMATION:
APPLICANT: DMITROVSKY, ETHAN
APPLICANT: WARRELL JR, RAYMOND P
APPLICANT: MILLER JR, WILSON H
APPLICANT: FRANKEL, STANLEY
TITLE OF INVENTION: METHODS FOR THE DETECTION AND
TREATMENT OF ACUTE PROMYELOCYTIC LEUKEMIA (APL)
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DINIUM LLP
STREET: 1185 AVENUE OF THE AMERICANS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,728B
FILING DATE: 21-JUL-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/673,838
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 38694-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391 0525
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2928 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: hRAR ALPHA
FEATURE:
NAME/KEY: CDS
LOCATION: 103..1488
OTHER INFORMATION:
US-08-095-728B-3
Query Match 5.9%, Score 74.4; DB 2; Length 2928;
Best Local Similarity 60.3%; Pred. No. 2, 90-10;
Matches 123; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
DB 286 TGGCTGTGACGATTAAGCATGAGATATGAGAGTGTGTGTGAGAGA 345
DB 364 TGGTTGTGTAAGAAATGCTGAACTAGTATGAGAGTGTGTGTGAGAGA 424
DB 346 TGAAGGCTTTTAAAGAGCATTAAGGACATATGATATTTGCGAGTACA 405
DB 424 TGAAGGCTTTTAAAGAGCATTAAGGACATATGATATTTGCGAGTACA 483
DB 406 AATCAGTGTACATATTAATAAAGAGGAGTGTGTGTGAGAGTGTGTGAG 465
DB 484 AATCAGTGTACATATTAATAAAGAGGAGTGTGTGTGAGAGTGTGTGAG 544
DB 466 TGTTCGAGAGTGTGTGTGAGAGTGTGTGTGAGAGTGTGTGTGAGAG 489
DB 544 TGTTCGAGAGTGTGTGTGAGAGTGTGTGTGAGAGTGTGTGTGAGAG 567

RESULT 33
PCT-US92-02320A-3
Sequence 3, Application PCT/US9202320A
GENERAL INFORMATION:
APPLICANT: Sloan Kettering Institute, for Cancer Research
TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER

```

1  NUMBER OF SEQUENCES: 6
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: COOPER & DUNHAM
4  STREET: 40 ROCKEFELLER PLAZA
5  CITY: NEW YORK
6  STATE: NEW YORK
7  COUNTRY: U.S.A.
8  ZIP: 10112
9  COMPUTER READABLE FORM:
10 MEDIUM TYPE: floppy disk
11 COMPUTER: IBM pc compatible
12 OPERATING SYSTEM: pc-dos/MS-DOS
13 SOFTWARE: Patient In Release #1.24
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: PCT/US92/02120A
16 FILING DATE:
17 CLASSIFICATION:
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 6,75,084
20 FILING DATE: 22 MAR 1991
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 6,75,838
23 FILING DATE: 22 MAR 1991
24 ATTORNEY/AGENT INFORMATION:
25 NAME: WHITE, JOHN P.
26 REGISTRATION NUMBER: 28,678
27 REFERENCE/DOCKET NUMBER: 68694-PCT
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (212) 977 9550
30 TELEFAX: (212) 644 0525
31 TELETYPE: (212) 422523 (Cable)
32 INFORMATION FOR SEQ ID NO: 1:
33 LENGTH: 2928 base pairs
34 STRANDNESS: double
35 TYPE: nucleic acid
36 TOPOLOGY: linear
37 MOLECULE TYPE: cDNA
38 IMMEDIATE SOURCE:
39 CLONED FROM ALPHA
40 NAME/KEY: COS
41 LOCATION: 105, 1488
42 OTHER INFORMATION:
43 PCT US92 02120A
44
45 Query Match 5,983 Score 74.4 DB 52 Length 2928:
46 Best Local Similarity 60.8% Pctd. No. 2,90-103
47 Matches 1245 Conservative 0 Mismatches 81 Indels 0 Gaps 0
48
49 QY 286 TGGCTGCTGAGAGATTAAGATGAGGATATGATATGAGGCTGCTGAGAGCA 445
50 111 11111 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
51 DB 664 TGGCTGCTGAGAGATTAAGATGAGGATATGATATGAGGCTGCTGAGAGCA 424
52 111 11111 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
53 QY 446 TGTAAAGGCTTTTAAAGAAATGATTCAGAGCAATATGATATATTTGTCGATGCA 405
54 111 111 111 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
55 DB 424 TGTAAAGGCTTTTGGGCGGAGATTCAGAGCAATATGATATATTTGTCGATGCA 484
56 111 111 111 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
57 QY 406 AATACAGTGAATGATTAAGAAACCGGCGAAGATGCGAGGCTTGGAGCTTCGAG 465
58 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
59 DB 484 AATACAGTGAATGATTAAGAAACCGGCGAAGATGCGAGGCTTGGAGCTTCGAG 544
60 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
61 QY 466 TGTATGAAGTGGTAATGATGAG 489
62 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
63 DB 544 TGTATGAAGTGGTAATGATGAG 567
64 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
65
66 RESULT 34
67 US 08 592 483 1
68 Sequence 1, Application us-08-826-361a-2
69 Patent No. 5850760
70 GENERAL INFORMATION:
71 APPLICANT: Tsai, S. and S.J. Collins

```

```

1  TITLE OF INVENTION: "Hematopoietic Cell Lines Bearing Altered Retinoid Acid Re
2  NUMBER OF SEQUENCES: 10
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE: Christenson, O'Connor, Johnson and Kindness
5  STREET: 2800 Pacific First Centre, 1420 Fifth Avenue
6  CITY: Seattle
7  STATE: Washington
8  COUNTRY: USA
9  ZIP: 98101
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Diskette-5.25 inch, 1.2MB storage
12 COMPUTER: IBM PC/486 compatible
13 OPERATING SYSTEM: MS-DOS 4.01
14 SOFTWARE: Word for Windows 5.01-t
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/592, 183
17 FILING DATE:
18 CLASSIFICATION: 435
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: 08/099, 242
21 FILING DATE: July 28, 1994
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Broderick, Thomas F.
24 REGISTRATION NUMBER: 31,332
25 REFERENCE/DOCKET NUMBER: FHC-1-7190
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: 1-206-682 8100; 1-206-224 0709(direct)
28 INFORMATION FOR SEQ ID NO: 1:
29 LENGTH: 2940 base pairs
30 STRANDNESS: double
31 TYPE: nucleic acid
32 TOPOLOGY: linear
33 MOLECULE TYPE: cDNA to mRNA
34 DESCRIPTION: see Figure 16B; RAR-alpha
35 IMMEDIATE SOURCE:
36 LIBRARY: cDNA
37
38 US-08-592-383-1
39
40 Query Match 5,983 Score 74.4 DB 2: Length 2940:
41 Best Local Similarity 60.8% Pctd. No. 2,90-103
42 Matches 1245 Conservative 0 Mismatches 81 Indels 0 Gaps 0
43
44 QY 286 TGGCTGCTGAGAGATTAAGATGAGGATATGATATGAGGCTGCTGAGAGCA 445
45 111 11111 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
46 DB 664 TGGCTGCTGAGAGATTAAGATGAGGATATGATATGAGGCTGCTGAGAGCA 424
47 111 11111 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
48 QY 446 TGTAAAGGCTTTTAAAGAAATGATTCAGAGCAATATGATATATTTGTCGATGCA 405
49 111 111 111 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
50 DB 424 TGTAAAGGCTTTTGGGCGGAGATTCAGAGCAATATGATATATTTGTCGATGCA 484
51 111 111 111 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
52 QY 406 AATACAGTGAATGATTAAGAAACCGGCGAAGATGCGAGGCTTGGAGCTTCGAG 465
53 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
54 DB 484 AATACAGTGAATGATTAAGAAACCGGCGAAGATGCGAGGCTTGGAGCTTCGAG 544
55 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
56 QY 466 TGTATGAAGTGGTAATGATGAG 489
57 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
58 DB 544 TGTATGAAGTGGTAATGATGAG 567
59 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
60
61 RESULT 35
62 5171671 1
63 Patent No. 5171671
64 APPLICANT: EVANS, RONALD M.; JONG, ESTELITA S.; SEED, J.
65 PRONIMAR S.; THOMPSON, CATHERINE C.; THOMPSON, KAZUHIKO
66 FUCHIERE, VINCENT
67 TITLE OF INVENTION: RETINOIC ACID RECEPTOR COMPOSITION
68 NUMBER OF SEQUENCES: 2
69 CURRENT APPLICATION DATA:
70 APPLICATION NUMBER: 05/337,546, 256
71 FILING DATE: 06 AUG 1990
72 PRIOR APPLICATION DATA:
73 APPLICATION NUMBER: 276, 536

```

US-08-306-691B-52

NAME/KEY: QDS
LOCATION: 67.2457
OTHER INFORMATION:

US - 08 - 095 - 728B -

[illegible]

```

1  RESULT 49
2  US 08 412 312A 1/2
3  Sequence 1, Application US/08412312A
4  Patent No. 5550024
5  GENERAL INFORMATION:
6  APPLICANT: Rothschild, Max F.
7  APPLICANT: Tugale, Christopher K.
8  APPLICANT: Jacobson, Carol D.
9  APPLICANT: Mileham, Alan J.
10 APPLICANT: Plastow, Graham S.
11 APPLICANT: Vaske, David A.
12 TITLE OF INVENTION: Genetic Markers for Pig Litter Size
13 NUMBER OF SEQUENCES: 9
14 CORRESPONDENCE ADDRESS:
15 ADDRESSEE: Zarley, McKee, Thomey, Wooldress & Sease
16 STREET: 801 Grand Suite 3200
17 CITY: Des Moines
18 STATE: Iowa
19 COUNTRY: United States
20 ZIP: 50309
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: Floppy disk
23 COMPUTER: IBM PC compatible
24 OPERATING SYSTEM: PC DOS/MS-DOS
25 SOFTWARE: Patent In Release #1.0, Version #1.40
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: US/08/412,412A
28 FILING DATE: 26-SEP-1994
29 CLASSIFICATION: 435
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: US 07/961,819
32 FILING DATE: 19 OCT-1992
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: US 07/687,708
35 FILING DATE: 19-APR-1991
36 ATOMERGENCY/AGENT INFORMATION:
37 NAME: Nebel, Heidi S.
38 REGISTRATION NUMBER: 47,719
39 REFERENCE/CHECKED NUMBER: 4844 4
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: 515-288-4667
42 TELEFAX: 515-288-1348
43 INFORMATION FOR SEQ ID NO: 1:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 4460 base pairs
46 TYPE: nucleic acid
47 STRANDEDNESS: double
48 TOPOLOGY: linear
49 MOLECULE TYPE: DNA (genomic)
50 HYPOTHEetical: NO
51 ANTI-SENSE: NO

```

```

; INFORMATION FOR SEQ ID NO: 13:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 3511 base pairs
;     TYPE: nucleic acid

```

Search completed: June 16, 2003, 03:20:47
Job time : 61.9084 secs

•

•

? TITLE OF INVENTION: Lipids
 ? FILE REFERENCE: 24799
 ? CURRENT APPLICATION NUMBER: US/10/259,864
 ? CURRENT FILING DATE: 2002-09-30
 ? PRIOR APPLICATION NUMBER: US 60/425,178
 ? PRIOR FILING DATE: 2001-09-28
 ? NUMBER OF SEQ ID NOS: 6
 ? SOFTWARE: Patent In version 4.1
 ? SEQ ID NO: 4
 ? LENGTH: 3400
 ? TYPE: DNA
 ? ORGANISM: Rat/human chimera
 ? FEATURE:
 ? NAME/KEY: CDS
 ? LOCATION: (1)-(3400)
 ? OTHER INFORMATION:
 ? US-10-259,864-4

Query Match 19.0%; Score 248; DB 9; Length 3400;

Host Local Similarity 55.7%; Pred. No. 3,20-63;

Matches 554; Conservative 0; Mismatches 410; Indels 29; Gaps 4;

Oy 284 TCTGCGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 343
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2075 TCTGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2134
 Oy 444 GATCTAACCGCTTTTAAACAGACATTCAGACATTAATGATTAATTTGGTACCTA 403
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2135 GCTGCAACGATTTCTTAAACAGACATTCAGACATTAATGATTAATTTGGTACCTA 2194
 Oy 404 CAATTAAGTGTACATTAATTAACAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 463
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2195 CAATTAAGTGTACATTAATTAACAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 2254
 Oy 464 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 523
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2255 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 2314
 Oy 524 TCTGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 583
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2415 TCTGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2473
 Oy 584 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 643
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2474 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2533
 Oy 644 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 693
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2434 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2493
 Oy 694 -----GCTACGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 748
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2494 GCTACGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2553
 Oy 749 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 798
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2554 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 2613
 Oy 799 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 858
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2614 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2673
 Oy 859 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 918
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2674 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 2733
 Oy 919 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 978
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2734 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2793
 Oy 979 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 1038
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 2794 GTCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 2853

RESULT 22

US-09-853-033-3

? Sequence 3; Application US/09853033

? Patent No. US2002010006B1

? GENERAL INFORMATION:

? APPLICANT: CHAMBER, PIERRE

? TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION

? FILE REFERENCE: 065691/0222

? CURRENT APPLICATION NUMBER: US/09/853,033

? PRIOR FILING DATE: 2001-05-11

? PRIOR APPLICATION NUMBER: FR 00/12570

? NUMBER OF SEQ ID NOS: 14

? SOFTWARE: Patent In Ver. 2.1

? SEQ ID NO: 3

? LENGTH: 1983

? TYPE: DNA

? ORGANISM: Artificial sequence

? FEATURE:

? NAME/KEY: CDS

? OTHER INFORMATION: Description of Artificial Sequence: Chimeric sequence

? US-09-853-033-3

Query Match 17.5%; Score 218.4; DB 10; Length 1983;

Host Local Similarity 61.5%; Pred. No. 3,10-57;

Matches 386; Conservative 0; Mismatches 236; Indels 6; Gaps 2;

Oy 626 GCTTACGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 685
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1121 GCTTACGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 1180
 Oy 686 TC---AGCGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 742
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1181 GCTTACGCGGTGTCTGAGCAATTCAGATTCGGATATGAGTCTGGCTGTGAG 1240
 Oy 743 TGGCGACAGAGAGTGTGACATTCAGATTCGGATATGAGTCTGGCTGTGAG 802
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1241 TGGCGACAGAGAGTGTGACATTCAGATTCGGATATGAGTCTGGCTGTGAG 1300
 Oy 803 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 862
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1301 AGCTTACGACAGATTCAGACATTCAGACATTCGGATATGAGTCTGGCTGTGAG 1360
 Oy 863 TGGCGACAGAGAGTGTGACATTCAGATTCGGATATGAGTCTGGCTGTGAG 922
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1361 TGGCGACAGAGAGTGTGACATTCAGATTCGGATATGAGTCTGGCTGTGAG 1420
 Oy 923 TGGCGACAGAGAGTGTGACATTCAGATTCGGATATGAGTCTGGCTGTGAG 982
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1421 TGGCGACAGAGAGTGTGACATTCAGATTCGGATATGAGTCTGGCTGTGAG 1480

[illegible][illegible]

DB 1826 AAGCAGGATGACCTGCTGACCCGGAAGTAAAGAACTGGGAATCTGAATGCAAG 1885
541 GAGATATAGAAATTTCTTCAAGCAAGCTGACTGTGGGCAAGCAAGCAAGTGGGCG 530
DB 1886 AGCAAGATGAACTCTCAATCTCAAGCTCAAGCTCAAGCAAGCAAGCTCAAG 1445
591 GACGAGGAGGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 630
DB 1946 TGTCAACATTTGAAGGCTATAGTGTG-----TCAGCCATCTTTCTGAATGTC 1992
651 CACCTCTTCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 710
DB 1993 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 2051
711 GAGAGTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 770
DB 2052 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 2111
771 GAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 830
DB 2112 CAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 2171
831 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 884
DB 2172 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 2231
885 TCAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 926
DB 2232 TGTCAACATTTGAAGGCTATAGTGTG-----TCAGCCATCTTTCTGAATGTC 2273

RESULT 38

US-09-997-267-1
Sequence 1: Application US/09997267
Patent No. US20020165481A1
GENERAL INFORMATION:
APPLICANT: AURENS-FAITH, ISABELLE
APPLICANT: HAENDLER, BERNARD
TITLE OF INVENTION: HUMAN ANDROGEN RECEPTOR VARIANTS
FILE REFERENCE: SCH-1793
CURRENT APPLICATION NUMBER: US 09/997 267
PRIORITY FILING DATE: 2001-11-30
PRIORITY APPLICATION NUMBER: 60/255,078
PRIORITY FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 1
LENGTH: 1326
TYPE: DNA
ORGANISM: Homo sapiens
US-09-997-267-1

Query Match 7.6% Score 95.2% DB 9% Length 1326
Best Local Similarity 51.2% Pred. No. 5,80-192
Matches 332: Conservative 0: Mismatches 298: Indels 18: Gaps 4:
DB 243 GTGCTGATGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 302
445 ATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 404
DB 403 GTGCAAGGCTCTTCAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 362
405 AATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 464
DB 463 AATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 422
465 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 524
DB 423 AGTATATAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 482

QY 525 TGTGAGGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 584
DB 484 ACTATAGG-CAGCAAG 541
QY 545 TGGGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 644
DB 542 AGCTGAGGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 591
QY 645 AGCTGAGGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 704
DB 592 AATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 650
QY 705 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 764
DB 651 CACCTCTTCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 710
QY 765 GATGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 824
DB 711 GTGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 770
QY 825 GTGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 878
DB 771 GCTGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACT 830
QY 879 GTCAATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 926
DB 831 CACCAATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 878

RESULT 39

US-09-880-107-2300
Sequence 2: Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, David L.
APPLICANT: Voelkel, Joseph G.
APPLICANT: Scherf, Dwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
PRIORITY FILING DATE: 2001-06-14
PRIORITY APPLICATION NUMBER: 60/211,379
PRIORITY FILING DATE: 2000-06-14
PRIORITY APPLICATION NUMBER: US 60/237,054
PRIORITY FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO: 2300
LENGTH: 3715
TYPE: DNA
ORGANISM: Homo sapiens
OTHER INFORMATION: Genebank Accession No. US20020142981A1 M23263
US-09-880-107-2300

Query Match 7.6% Score 95.2% DB 10% Length 3715
Best Local Similarity 51.2% Pred. No. 1e-18
Matches 332: Conservative 0: Mismatches 298: Indels 18: Gaps 4:
QY 285 GTGCTGATGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 444
DB 2202 GTGCTGATGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 2261
QY 445 ATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 404
DB 2262 GTGCTGATGCTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 2321
QY 405 AATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 464
DB 2322 AATGTAAGGCTTTTAAAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 2381
QY 465 GTTAAAGGATGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGTAACTGAGT 524

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo
 1 (bases 1 to 947)
 NIH_MGC http://www.ncbi.nlm.nih.gov/
 AUTHORS
 TITLE
 JOURNAL
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: rga@bbsr@mail.nih.gov
 Tissue procurement: ATCC
 cDNA library preparation: Rubin laboratory
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILM)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILM at:
 http://image.llnl.gov
 Plate: L10M1998 row: h column: 09
 High quality sequence stop: 608.
 Location/Qualifiers
 1..937
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5478584"
 /clone_lib="NIH_MGC_98"
 /issue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Brain; Vector: pGB19; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming, directional
 cloning into EcoRI/XhoI sites using the following 5'-
 adaptor: ATGAGAGAGGCTT library constructed by Fred Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library." 2 others.

FEATURES SOURCE

Query Match 8.18; Score 101.4; DB 13; Length 937;
 Best local similarity 47.9%; Pred. No. 6,4e-18;
 Matches 291; Conservative 0; Mismatches 316; Indels 0; Gaps 0;

Y 334 TCTGTGAGGAGATGTAAGGCTTTTAAAGACAGATTCAGCATATGATTAAT 393
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 5 TCTGTGAGGCTTTTAAAGGCTTTTAAAGGAGATTCAGCATATGATTAAT 64
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 394 TCTGAGCTACAAATAGTGTACATATGATTAAGGAGATTCAGCATATGATTAAT 453
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 65 TGTGGGCTTCATATAGTGTGATGATGACAGGAGAGAGAGAGAGAGAGAGAG 124
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 454 GAACTTGGATATGATTAAGTATGATGATGATGATGATGATGATGATGATGAT 513
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 125 GCTGTGACCAAGTGTGAGGAGTGTGATGATGATGATGATGATGATGATGATGAT 184
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 514 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 185 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 244
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 574 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 633
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 245 GCTTGGGCTGTGAG 304
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 634 GCTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 693
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 305 AATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 694 GCTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 753
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 365 GCTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 754 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 813
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 425 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 814 TTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 873
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 485 TTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 874 TCTGAGCTACAAATAGTGTACATATGATTAAGGAGATTCAGCATATGATTAAT 933
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 545 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 934 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 993
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 605 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 23
 B058831 568 bp mRNA linear EST 11-DEC-2001
 B058831 NIBB Mochii normalized Xenopus tailbud library Xenopus
 laevis cDNA clone X1060113 5', mRNA sequence.
 ACCESSION
 B058831 GI:17492544
 VERSION
 B058831.1 GI:17492544
 KEYWORDS
 EST
 SOURCE
 African clawed frog.
 ORGANISM
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Anura; Anura; Mesobatrachia; Pipridae; Epidelae;
 Xenophelia; Xenopus.
 1 (bases 1 to 558)
 Kitayama, A., Torasaka, C., Mochii, M., Ueno, N., Shio, T., and Kohata
 Y.

TITLE
 JOURNAL
 COMMENT
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadashi Shio-1
 Center For Genetic Resource Information
 National Institute of Genetics
 111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshio@genes.nig.ac.jp.
 Note: This is a NIG_Mochii library."

FEATURES
 SOURCE
 1..558
 /organism="Xenopus laevis"
 /db_xref="taxon:8495"
 /clone_image="X1060113"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"
 /issue_type="whole embryo"
 /dev_stage="stage 25"
 /note="Vector: pBSK3; Site_1: NotI; Site_2: EcoRI; cDNAs
 were oligo-dT primed and directionally cloned. Cloning
 according to Mieskeop and Huber. Library is subcloned
 and was constructed by N. Garrett and A.M. Zorn,
 (Wellcome/CRC Institute)."

Query Match 8.08; Score 100.4; DB 13; Length 558;
 Best local similarity 51.6%; Pred. No. 9.3e-18;
 Matches 292; Conservative 0; Mismatches 256; Indels 18; Gaps 0;

Y 307 GATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 366
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 1 GATTAAGGCTTTTGGAG 60
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 457 AATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 41 AATTAAGGCTTTTGGAG 120
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 427 AATTAAGGCTTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 486
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 121 AATTAAGGCTTTTGGAG 180
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Y 487 AATTAAGGCTTTTGGAG 546
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 D 181 AATTAAGGCTTTTGGAG 240
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Downloaded from <http://ajphaphysocpharm.sagepub.com/> at 11:01 11 November 2014

Journal Pre-proof

[illegible][illegible]

REFERENCE	AUTHORS	TITLE
Mammalia; Insectaria; Crocodylia; Chelonia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Catarrhini; Hominoidea; Homo, 1 (bases 1 to 349)	Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R., Nadal,M.A., da Silva W Jr., Zardim,A., Pereira,Costa F.F.	
Gedali,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,V.V., O'Hare M.J., Soares,F., Brantani,R., Reis,I.F., de Souza,S.I. and Simpson,A.J.		
Shoiguun sequencing of the human transcriptome with cDNA expressed sequence tags		
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
Contact : Simpson A.J.G.		

Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, São Paulo-SP,
Brazil
Tel: +55-11-27044922
Fax: +55-11-2707001
Email: astimpson@ludwig.org.br
This sequence was derived from the FAPSP/IGR Human Cancer Genome
Project. This entry can be seen in the following URL:
http://www.ludwig.org.br/IGR/hc/chr11/chr11_0000024_F01_0000024
061200.024 chr11:0000024:12-06544 1)
Seq primer: pos 18 forward
High quality sequence stop: 70.
Location/Qualifiers
1..399

/alt_xref "taxon:3636"
 /clone_lib "GN0194"
 /dev_stage "Adult"
 /host="organ: Placenta normal; Vector: puc18; Site_1: Sma
 ; Site_2: Sma; A mini-library was made by cloning
 products derived from ORSIES PCR (U.S. Letters Patent
 Application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the puc18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions. "

BASE COUNT	ORIGIN
107 a	86 c 111 g 95 t

Query Match	7-38	Score 91	DB 12	Length 399
Host Local Similarity	63.58	Pred. No.	4.1e-15	
MisMatches	199	Conservation	0	MisMatches 80
				Indels 0
				Gaps 0
QY	292	GTCTGCAAGCATTAATGCAATGCGGCAATGAGACTCTGGCTGTCTGAAGAGATGAAG	351	
Ph	138	CTCTGTCTGCAATATGCTCTTCTGCTCTTCTGCTCTCTGCAATGATCTGAAGGCTCTGAAG	197	
QY	352	GGCTTTTAAAGAGAGGATTCAGAGGAGATTAAGATTAATTTGTGACATGCAAAATGAG	411	
Ph	198	CTATTCTTAAAGAGCAATTTCAAGGGCAATATGCAATATGCTGTGCTCTGCTCTGCAATGAA	257	
QY	412	TGTACATATGATAAAAACTGCGCGAAGAGCTGTGCAAGGCTGGCGATTCGGAAGATGTAAC	471	
Ph	258	TGTGAATATCAAAAAGCGACAGCGATAAAGCTGAGAGAGCTGTGATATGAAAGCTTTA	317	
QY	472	GAGCTCTGAATGCTGCAATGCTGGGCTCTCGCGAGAGCAAGA	510	
Ph	318	AAAGTGGGATCTGTAAAGAAAGAGGAGTGGCTTTGAACA	356	

RESULT 37	592 bp	mRNA	linear	EST 15-MAY-2001
RG746447				
LOCUS	60276372871	NC_0215	homo sapiens cDNA clone IMAGE:4856886	5
DEFINITION	mRNA sequence.			
ACCESSION	RG746447			
VERSION	RG746447.1	GI:14057100		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	homo sapiens			

REFERENCE	ARTICLES	TITLE	JOURNAL	COMMENT
1 (bases 1 to 592)		Mammalia: Eutheria: Primates) Catarrhini: Hominoidea: Homo.		
NIH-MG	http://imga.nci.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Published (1999)	
Contact: Robert Strausberg, Ph.D.		Email: cgabs-r@mail.nih.gov		
Tissue Procurement: ATCC		cDNA Library Preparation: Ling Hong/Rubin Laboratory		
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMI)		DNA Sequencing by: NIH Intramural Sequencing Center		
Clone distribution: MGC clone distribution information can be		found through the I.M.A.G.E. Consortium/ILMI at:		
http://imga.imgl.gov		Plate: ILMI1708	row: h	column: 07
High quality sequence stop: 587.				

```

source
1. 1992
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cdo="TMAGP_4856886"
/clone_lib="NH MGC 15"
/tissue_type="adipocytic tumor cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: colon; Vector: pMT7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GCGAAGAG(G). Size selected <500bp for average
insert size 1.6kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using GAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 98 A 182 C 197 G 115 T
ORIGIN
CCTGTCATGTCGACGACATTCAGCGTGGCGATATCACTATGAGAGTCGTCGCTGTCGAGG 344

```


TITLE
TSUMODA, Y., WATAHIKI, A., WATANABE, S., YAMAMURA, T., YAMANO, I.,
YANO, F., YASUNISHI, A., YOKOTA, I., YOSHIDA, K., YOSHIDA, A., YOSHINO
M., MURAMATSU, M. and HAYASHIZAKI, Y.
RIKEN MOUSE ESTS (Kono, H., et al.)

COMMENT
Contact. Yoshitake Hayashizaki
Institute for Genome Functional Research, Graduate School of Science,
University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Laboratory for genome exploration research group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-33 Esaka, Wako-shi, Saitama Prefecture 350-0045, Japan

Tel: 81-45-503-9222

EX: 81 45 509 2.16

Ensidia : [genome-ics.riken.go.jp/](http://genome-ics.ics.riken.go.jp/)
 IPI : <http://genome-ics.riken.go.jp/>

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, M., Okazaki, Y., Muramatsu, M., and Hasegawa, R. Y.

Thermostabilization and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length
CDM Proc Natl Acad Sci U S A 95(73): 520-524 (1998)

Itch, M., Kitsuunai, T., Akiyama, J., Shibata, K., Iwawa, M., Kawai, J.,
DNA. Proc. Natl. Acad. Sci. U.S.A. 75 (2), 220-224 (1978)

Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hanaoka, K. 2003. The human genome database: a model for genome-wide research. *Genomics* 81:1-10.

Automated filtration-based high-throughput plasmid preparation

system (genome pos 4 (5), 4n4-471) (1999)
Genomic and metabolic

High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 303.

19-44 (1999)

For further details, please visit our web site (<http://journals.oxfordjournals.org>) for

FEATURES	LOCATION/Qualifiers
COURSE	1 075

```

1. .275
source
/organism="Mus musculus"

```

```
/db_xref="taxon:10090"
xref="AA:G00440"
```

```
/clone_lib="RIKEN full-length enriched, adult male spinal
```

```
cord"
/cov="major"
```

```
/sex= male
/tissue_type="spinal cord"
```

```
/dev_stage="adult"
/lab_host="192.168"
```

/note="Site_1: SalI; Site_2: BamHI; cDNA library was

Prepared and Sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Project of genome exploration research group in seven
Genomic Sciences Center and Genome Science Laboratory in

RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer 15'

GAGAGGAGGAAGATCCAAAGAGCTCTTTT TTT TTT TTT TTT VN 3'), cDNA was prepared by using tobacco etch virus activated reverse

prepared by using creatase (Sigma) activated reverse transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of normalization to Pot + 10.0 and subtraction to Pot + 459.0. Second

co for 10.0 and subtraction co for = 499.0; second strand cDNA was prepared with the primer adapter of

sequence [5' GAGGAGAGGATTCTCGACTTAATTAATAATTCACCCCCCCCCCC

37]. cDNA was cloned with XbaI and BamHI vectors as

3. The plasmid was cleaved with *Xba*I and *Pst*I and the resulting fragments were ligated into the modified pRlucrescript KS(+) after bulk excision from Lambda

BASE COUNT	FLC I. "
75	65
68	67

PROJ	COONT	12	9	07
ORIGIN				

Answer Match 7 99. Score 89 6. DB 1/0. Length 375.

Test	Mean	SD	Min	Max
First Local Similarity	64.48	10.80	30-15	80-100

Matches	134;	Conservative	0;	Mismatches	74;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	-----	--------	----	------	----

276 TGGTAACTTTGGTGTATGAGTAACTATGCGATAATATAGCACTGCGTC 335

[illegible]

336 GTGTGAAGGATGTAAAGGCTTTTAAAGAGAGCAATTCACAGACATAATGATTATATTTG 395

114 G T G T T G A A C T T G T A A T T C T T T T T A A A A G A G C A G T G C A A T G A C A C A C A A T T A C C T T G 173

395 10'AA'1A'AAATAA111A'AA11'1A1AAAA'10'10'AAAGCTG'CAAGGCTGGG 455

2

DB 174 TCGTCGAACAAATGATTCATTCATTAATAATTCGAACAAAAATTCGCAATGATTCG 233

CY 456 ACTTCGGAATCTTACGAATCGGAATG 483

DB 234 CTATCGAAAAATCTTCAACTGGAAATG 261

Search completed: June 16, 2003, 04:16:07
 Job Time = 1906.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

1M protein - protein search, using sw model

Run on: June 16, 2003, 03:20:57 : Search time 125.704 seconds

(without alignments)
505,642 Million cpd1 updates/sec

Title: US-08-826-361a-5
Perfect score: 2522
Sequence: 1 MVSILPSNNTNLFKPGPGQT PPSAFPSKSPGSGNNGSD 477

Scoring table:
Matrix: BLOSUM62
Gap: 10.0, Gapext: 0.5

Searched: 908470 seqs, 13250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 120 summaries

```
Database :
1: A.Geneseq_101002:*
2: /SID82/gcdata/geneseq/geneseq-emb1/AA1980.PAT*
3: /SID82/gcdata/geneseq/geneseq-emb1/AA1981.PAT*
4: /SID82/gcdata/geneseq/geneseq-emb1/AA1982.PAT*
5: /SID82/gcdata/geneseq/geneseq-emb1/AA1983.PAT*
6: /SID82/gcdata/geneseq/geneseq-emb1/AA1984.PAT*
7: /SID82/gcdata/geneseq/geneseq-emb1/AA1985.PAT*
8: /SID82/gcdata/geneseq/geneseq-emb1/AA1986.PAT*
9: /SID82/gcdata/geneseq/geneseq-emb1/AA1987.PAT*
10: /SID82/gcdata/geneseq/geneseq-emb1/AA1988.PAT*
11: /SID82/gcdata/geneseq/geneseq-emb1/AA1989.PAT*
12: /SID82/gcdata/geneseq/geneseq-emb1/AA1990.PAT*
13: /SID82/gcdata/geneseq/geneseq-emb1/AA1991.PAT*
14: /SID82/gcdata/geneseq/geneseq-emb1/AA1992.PAT*
15: /SID82/gcdata/geneseq/geneseq-emb1/AA1993.PAT*
16: /SID82/gcdata/geneseq/geneseq-emb1/AA1994.PAT*
17: /SID82/gcdata/geneseq/geneseq-emb1/AA1995.PAT*
18: /SID82/gcdata/geneseq/geneseq-emb1/AA1996.PAT*
19: /SID82/gcdata/geneseq/geneseq-emb1/AA1997.PAT*
20: /SID82/gcdata/geneseq/geneseq-emb1/AA1998.PAT*
21: /SID82/gcdata/geneseq/geneseq-emb1/AA1999.PAT*
22: /SID82/gcdata/geneseq/geneseq-emb1/AA2000.PAT*
23: /SID82/gcdata/geneseq/geneseq-emb1/AA2001.PAT*
24: /SID82/gcdata/geneseq/geneseq-emb1/AA2002.PAT*
```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No	Score	Query Match	Length	DB ID	Description
1	2522	100.0	477	18	AAW14724
2	2522	100.0	477	20	AAW01597
3	2522	100.0	477	23	AAW04436
4	2522	100.0	477	24	AAW04437
5	2522	100.0	477	25	AAW04438
6	2522	100.0	477	26	AAW04439
7	2522	100.0	477	27	AAW04440
8	2522	100.0	477	28	AAW04441
9	2522	100.0	477	29	AAW04442
10	2522	100.0	477	30	AAW04443

11	2512	99.6	485	18	AAW14724	Human oestrogen re
12	2493	98.9	485	20	AAW01597	Human oestrogen re
13	2257	89.5	549	20	AAW04436	Human oestrogen re
14	2255	89.4	549	23	AAW04437	Human oestrogen re
15	2249	89.2	485	18	AAW14724	Human oestrogen re
16	2246	89.2	485	20	AAW01597	Human oestrogen re
17	2246	89.2	485	23	AAW04436	Human oestrogen re
18	2231.5	88.5	474	21	AAW1946	Human oestrogen re
19	2230	88.4	485	18	AAW14724	Human oestrogen re
20	2212	87.2	416	18	AAW04433	Human oestrogen re
21	2198	87.2	416	18	AAW04434	Human oestrogen re
22	2198	87.2	416	18	AAW04435	Human oestrogen re
23	2198	87.2	416	18	AAW04436	Human oestrogen re
24	2198	87.2	416	18	AAW04437	Human oestrogen re
25	2000.5	79.3	510	20	AAW04438	Human oestrogen re
26	2000.5	79.3	510	20	AAW04439	Human oestrogen re
27	1499	59.4	474	20	AAW04440	Human oestrogen re
28	1499	59.4	474	20	AAW04441	Human oestrogen re
29	1211	48.0	587	23	AAW04436	Human oestrogen re
30	1210	48.0	587	23	AAW04437	Human oestrogen re
31	1209.5	48.0	587	23	AAW04438	Human oestrogen re
32	1206	47.8	595	22	AAW04439	Human oestrogen re
33	1204	47.7	595	22	AAW04440	Human oestrogen re
34	1204	47.7	595	22	AAW04441	Human oestrogen re
35	1204	47.7	595	22	AAW04442	Human oestrogen re
36	1204	47.7	595	22	AAW04443	Human oestrogen re
37	1204	47.7	595	22	AAW04444	Human oestrogen re
38	1202	47.7	595	22	AAW04445	Human oestrogen re
39	1201	47.6	595	22	AAW04446	Human oestrogen re
40	1197	47.5	595	22	AAW04447	Human oestrogen re
41	1197	47.5	595	22	AAW04448	Human oestrogen re
42	1195	47.4	595	22	AAW04449	Human oestrogen re
43	1195	47.4	595	22	AAW04450	Human oestrogen re
44	1195	47.4	595	22	AAW04451	Human oestrogen re
45	1195	47.4	595	22	AAW04452	Human oestrogen re
46	1195	47.4	595	22	AAW04453	Human oestrogen re
47	1195	47.4	595	22	AAW04454	Human oestrogen re
48	1195	47.4	595	22	AAW04455	Human oestrogen re
49	1195	47.4	595	22	AAW04456	Human oestrogen re
50	1187	47.1	595	22	AAW04457	Human oestrogen re
51	1187	47.1	595	22	AAW04458	Human oestrogen re
52	1186.5	47.0	595	22	AAW04459	Human oestrogen re
53	1181	46.8	595	22	AAW04460	Human oestrogen re
54	1181	46.8	595	22	AAW04461	Human oestrogen re
55	1181	46.8	595	22	AAW04462	Human oestrogen re
56	1181	46.8	595	22	AAW04463	Human oestrogen re
57	1176	46.5	595	22	AAW04464	Human oestrogen re
58	1176	46.5	595	22	AAW04465	Human oestrogen re
59	1176	46.5	595	22	AAW04466	Human oestrogen re
60	1176	46.5	595	22	AAW04467	Human oestrogen re
61	1176	46.5	595	22	AAW04468	Human oestrogen re
62	1176	46.5	595	22	AAW04469	Human oestrogen re
63	1176	46.5	595	22	AAW04470	Human oestrogen re
64	1176	46.5	595	22	AAW04471	Human oestrogen re
65	1176	46.5	595	22	AAW04472	Human oestrogen re
66	1176	46.5	595	22	AAW04473	Human oestrogen re
67	1176	46.5	595	22	AAW04474	Human oestrogen re
68	1176	46.5	595	22	AAW04475	Human oestrogen re
69	1176	46.5	595	22	AAW04476	Human oestrogen re
70	1176	46.5	595	22	AAW04477	Human oestrogen re
71	1176	46.5	595	22	AAW04478	Human oestrogen re
72	1176	46.5	595	22	AAW04479	Human oestrogen re
73	1176	46.5	595	22	AAW04480	Human oestrogen re
74	1176	46.5	595	22	AAW04481	Human oestrogen re
75	1176	46.5	595	22	AAW04482	Human oestrogen re
76	1176	46.5	595	22	AAW04483	Human oestrogen re
77	1176	46.5	595	22	AAW04484	Human oestrogen re
78	1176	46.5	595	22	AAW04485	Human oestrogen re
79	1176	46.5	595	22	AAW04486	Human oestrogen re
80	1176	46.5	595	22	AAW04487	Human oestrogen re
81	1176	46.5	595	22	AAW04488	Human oestrogen re
82	1176	46.5	595	22	AAW04489	Human oestrogen re
83	1176	46.5	595	22	AAW04490	Human oestrogen re

84	644.5	25.5	4.98	2.1	AAV8.082.3	Human nucleofor receptor
85	643	25.5	4.98	2.0	AAV2.946.7	Human SUPER84 protein
86	643	25.1	4.83	2.1	AAV1.297.0	Human osteopontin
87	640	25.0	4.33	9	AAV0.909.1	Sequence of human
88	640	25.0	4.33	2.1	AAV0.999.6	Human ERBB2 a protein
89	640	25.0	5.00	2.0	AAV0.384.6	Human nucleofor receptor
90	640	25.0	5.00	2.1	AAV8.082.2	Human nucleofor receptor
91	598.5	24.7	3.74	2.2	AAV0.998.6	Antidiotypic antibody
92	598.5	24.7	3.82	2.2	AAV0.998.6	Antidiotypic antibody
93	597	24.7	4.86	2.1	ABH05.347	Human nucleofor receptor
94	597	24.7	5.52	2.2	AAV0.941.1	Human polypeptide
95	594.5	24.6	5.21	9	AAV0.909.1	Sequence of human
96	588.5	24.1	5.56	2.2	AAV4.11.9.2	Human polypeptide
97	582.5	22.7	4.18	2.1	AAV8.082.4	Human nucleofor receptor
98	570.5	22.6	4.18	2.0	AAV0.384.6	Human nucleofor receptor
99	522	20.7	4.84	2.2	ABH6.325.1	Proteoglycan molecule
100	504	19.0	9.9	2.2	AAV3.71.7.3	Peptide #11210 core
101	460.5	18.3	5.25	2.0	AAV2.162.5	Ligand binding domain
102	458.5	18.2	5.33	14	AAV3.046.8	BRXK1-2011 - Human
103	458.5	18.2	5.33	16	AAV2.248.3	Human H-2RIIIP - Human
104	458.5	18.2	5.33	2.3	AAV0.782.9.7	Human Rct-101d X R
105	455.5	18.1	4.46	1.2	AAV1.118.3.7	Human Rct-101d X R
106	455.5	18.1	4.46	1.5	AAV2.748.4	H-2RIIIP - Nucleofor
107	453.5	18.0	4.48	1.4	AAV3.046.7	BRXK1-2011 - Human
108	453.5	18.0	5.10	1.4	AAV3.046.9	BRXK1-2011 - Human
109	448	17.8	9.33	2.0	AAV2.162.1	Ligand binding domain
110	448	17.8	9.33	2.1	AAV2.297.7	Human prostatic acid
111	443	17.6	4.63	1.2	AAV1.080.0	Mouse RKR-gamma toxin
112	443	17.6	4.63	1.4	AAV1.047.1	Mouse RKR-gamma toxin
113	440.5	17.5	9.18	1.2	AAV1.222.3	Human androgen receptor
114	440	17.4	4.63	2.3	AAV0.829.8	Human Rct-101d X R
115	439.5	17.4	3.59	2.3	AAV6.882.8	Fused androgen receptor
116	437.5	17.3	4.52	2.0	AAV2.162.7	Ligand binding domain
117	437.5	17.3	4.18	2.0	AAV3.046.1	Human androgen receptor
118	437.5	17.3	9.19	1.0	AAV3.109.9	Human androgen receptor
119	437.5	17.3	9.19	1.0	AAV3.099.6	Human androgen receptor
120	437.5	17.3	9.19	1.8	AAV4.788.4	Androgen receptor

ALLEGMENTS

XX	RESEPT 1
XX	AAM3421Z
ID	AAM3421Z standard; Proteinc; 477 AA.
XX	
Ac	AAM3421Z;
XX	
Dt	20 APR 1998 (first entry)
DE	Human estrogen receptor protein.
XX	
KW	estrogen receptor; steroid; alternative splicing; estradiol;
KM	estrogen; estradiol; screening.
XX	
OS	Homo Sapiens.
XX	
PB	EPI798478-A2.
XX	
PD	01 OCT 1997.
XX	
FE	25 MAR 1997; 97EP-020904.
XX	
FE	22 NOV 1996; 96EP-020420A.
PK	26 MAR 1996; 96EP-020582D.
XX	
FA	(ALKE) AKZO NOBEL NV.
FA	
FV	[1] Kema R., Mossmann St,
XX	
XX	WET: 1997-474188/A4.
DB	N P5006; AAT88412.
XX	

Query Match	100.0%	Score 2522	DB 18	Length 477
blast Local Similarity	100.0%	Ident. No. 2	1e-291	
Matches 477	Conservative 0	Mismatches 0	Indels 0	Gaps 0
01	MNSYSINVTNLEGGCHGPOTSPNVLWPTQCHLSPIYVHQLSHLKAEPQSKHFEAST	60		
02	1 MNSYSINVTNLEGGCHGPOTSPNVLWPTQCHLSPIYVHQLSHLKAEPQSKHFEAST	60		
03	1 MNSYSINVTNLEGGCHGPOTSPNVLWPTQCHLSPIYVHQLSHLKAEPQSKHFEAST	60		
04	61 FHTLVNRETLKKKSGNRCASVPTGSKKDAHFVAVSDVASGYNGVASTETQKAF	120		
05	61 FHTLVNRETLKKKSGNRCASVPTGSKKDAHFVAVSDVASGYNGVASTETQKAF	120		
06	61 FHTLVNRETLKKKSGNRCASVPTGSKKDAHFVAVSDVASGYNGVASTETQKAF	120		
07	121 KPSLGGHNYVTPATNPTLTKNPKSCDAFLKLYEGVMKQSGRRECVKLVKQK	180		
08	121 KPSLGGHNYVTPATNPTLTKNPKSCDAFLKLYEGVMKQSGRRECVKLVKQK	180		
09	121 KPSLGGHNYVTPATNPTLTKNPKSCDAFLKLYEGVMKQSGRRECVKLVKQK	180		
10	181 SADEQLHCAKAKAPSSDAFVPEFLTDAI SPED VTLLEAFPPVITSPDAETPAS	240		
11	181 SADEQLHCAKAKAPSSDAFVPEFLTDAI SPED VTLLEAFPPVITSPDAETPAS	240		
12	181 SADEQLHCAKAKAPSSDAFVPEFLTDAI SPED VTLLEAFPPVITSPDAETPAS	240		
13	241 MMSSTIKLAKELVHMSIAKKTQFVELSLFDVKKLLESGMREVLDMGIMKESIDHKK	300		
14	241 MMSSTIKLAKELVHMSIAKKTQFVELSLFDVKKLLESGMREVLDMGIMKESIDHKK	300		
15	301 LIPAPRLVDIDRQKQVEGCLLEFPMILATTSPREELKQHKRYLVKAMILLNSMYL	360		
16	301 LIPAPRLVDIDRQKQVEGCLLEFPMILATTSPREELKQHKRYLVKAMILLNSMYL	360		
17	361 VTATGVALDSKRIAHLLANVTALVWVIAKSGTSSGVSNGPLANMLLSHYHAKSKRM	420		
18	361 VTATGVALDSKRIAHLLANVTALVWVIAKSGTSSGVSNGPLANMLLSHYHAKSKRM	420		
19	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
20	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
21	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
22	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
23	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
24	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
25	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
26	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
27	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
28	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
29	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
30	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
31	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
32	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
33	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
34	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
35	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
36	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
37	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
38	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
39	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
40	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
41	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
42	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
43	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
44	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
45	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
46	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
47	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
48	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
49	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
50	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
51	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
52	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
53	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
54	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
55	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		
56	421 EHLINKKKNVNVVYVLLLEMLNAVLVGGTSSSTLSSGSSSADELSKSKESLQNSQ	477		

PR 25-MAR-1996: 96EP-0200820.
 PR 22-NOV-1996: 96EP-0203284.
 PR 25-MAR-1997: 97EP-0200903.
 XX
 XX (ALKU) AKZO NOBEL NV.
 XX
 P1 Mosselman S, Dijkema K;
 XX WPI: 2002-084414/12.
 OR N-PSDB: AA172128.
 XX
 PR New isolated chimeric receptor comprising a DNA binding domain and/or
 PT ligand binding domain of a new estrogen receptor, for identifying
 PT functional ligands or hormonal analogs for the receptor
 XX
 PS Example A: Page 15-17; 35pp; English.
 XX
 CC This sequence shows a novel estrogen receptor (ER). The gene encoding
 CC this new ER is located on chromosome 14 and has a different tissue
 CC distribution from classical ER. This ER also has two orphan ER's,
 CC ER-alpha and ER-beta. These orphan receptors have estrogen receptor
 CC related structure but do not appear to be able to bind estradiol or other
 CC ER ligands. The DNA binding domain (DBD) and ligand binding domain (LBD)
 CC from this ER may be used in the chimeric receptor of the invention.
 CC which also has an N-terminal domain. The chimeric receptor, or DNA
 CC encoding it, is useful in a screening assay for identification of new
 CC drugs. Similar chimeric receptors comprising the LBD of the new ER,
 CC and also comprising the DBD and an N-terminal domain derived from
 CC another nuclear receptor (e.g., progesterone receptor), are useful for the
 CC screening of compounds to identify new ligands or hormone analogs which
 CC are able to activate the new ER. Chimeric receptors comprising a DBD
 CC of the new ER, and LBD and an N-terminal domain derived from
 CC another nuclear receptor, can be used to identify new ligands or
 CC hormone analogs for the nuclear receptors.
 XX
 SV Sequence 477 AA:

Query Match 100.0%; Score 2522; DB 23; Length 477.
 Best Local Similarity 100.0%; Pred. No. 2; Ie-231;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 QY 1 MNYSIPSVNLTLEGGPGROTTSPNVLPPTGHLSPVVRHQLSHLYAFPKSWCFARSL 60
 DB 1 MNYSIPSVNLTLEGGPGROTTSPNVLPPTGHLSPVVRHQLSHLYAFPKSWCFARSL 60
 QY 61 EHTLPVNRETLKRKVSNGRCASPVTPGSKRDHAFCAVCSDVASGYHYGWSCEGKAF 120
 DB 61 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDHAFCAVCSDVASGYHYGWSCEGKAF 120
 QY 121 KRSTGHNVTGCPATNCTTIKKNPKSCQACPIKCYEVGMVCGSPREPGYPIVLRQR 180
 DB 121 KRSTGHNVTGCPATNCTTIKKNPKSCQACPIKCYEVGMVCGSPREPGYPIVLRQR 180
 QY 181 SADFQIHCAGAKKSGSHAPVRRLDLDALSPQIVLTLEAPRVLISRPAPRTFAS 240
 DB 181 SADFQIHCAGAKKSGSHAPVRRLDLDALSPQIVLTLEAPRVLISRPAPRTFAS 240
 QY 241 MNMSITKLADKELVHMLISMAKIPGEVELSLFDVRLLESCMMEVIMMLIMKSIHPCK 300
 DB 241 MNMSITKLADKELVHMLISMAKIPGEVELSLFDVRLLESCMMEVIMMLIMKSIHPCK 300
 QY 301 LIFAPDVLVDDEKCKVHGILIEFDMLATTISPRRELKIQHKRYLVKAMILLNSMYPL 360
 DB 301 LIFAPDVLVDDEKCKVHGILIEFDMLATTISPRRELKIQHKRYLVKAMILLNSMYPL 360
 QY 361 VTATQDADSSKRLHLNAVTDAIWWIAKSGISSQDSMBRLANLMLSHVHANSKCK 420
 DB 361 VTATQDADSSKRLHLNAVTDAIWWIAKSGISSQDSMBRLANLMLSHVHANSKCK 420
 QY 421 EHLNMGCKRVVVPYPLLEFMNAHVLESCSSITQSESPALSSKESGQNSQ 477
 DB 421 EHLNMGCKRVVVPYPLLEFMNAHVLESCSSITQSESPALSSKESGQNSQ 477

RESULT 3
 AAW33215
 ID AAW33215 standard: Protein: 530 AA.
 AC AAW33215:
 DI 20-APR-1998 (first entry)
 XX
 DE Human oestrogen receptor protein #2.
 KW Oestrogen receptor protein; steroid; alternative splicing; estradiol;
 KW estrogen; estradiol; screening.
 OS Homo sapiens.
 PN EP798478-A2.
 XX
 XX 01-OCT-1997.
 XX
 XX 25-MAR-1997: 97EP-0200903.
 PR 22-NOV-1996: 96EP-0203284.
 PR 25-MAR-1997: 96EP-0200820.
 XX
 XX (ALKU) AKZO NOBEL NV.
 P1 Dijkema K, Mosselman S;
 XX WPI: 1997-473188/44.
 OR N-PSDB: AA188415.
 XX
 PR DNA encoding estrogen receptor useful in screening assay to
 PT identify novel ligands or hormonal analogues
 XX
 PS Claim 4: Page 35-37; 45pp; English.

Query Match 100.0%; Score 2522; DB 18; Length 530.
 Best Local Similarity 100.0%; Pred. No. 2; Ie-231;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 QY 1 MNYSIPSVNLTLEGGPGROTTSPNVLPPTGHLSPVVRHQLSHLYAFPKSWCFARSL 60
 DB 54 MNYSIPSVNLTLEGGPGROTTSPNVLPPTGHLSPVVRHQLSHLYAFPKSWCFARSL 113
 QY 61 EHTLPVNRETLKRKVSNGRCASPVTPGSKRDHAFCAVCSDVASGYHYGWSCEGKAF 120
 DB 114 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDHAFCAVCSDVASGYHYGWSCEGKAF 173
 QY 121 KRSTGHNVTGCPATNCTTIKKNPKSCQACPIKCYEVGMVCGSPREPGYPIVLRQR 180
 DB 174 KRSTGHNVTGCPATNCTTIKKNPKSCQACPIKCYEVGMVCGSPREPGYPIVLRQR 233
 QY 181 SADFQIHCAGAKKSGSHAPVRRLDLDALSPQIVLTLEAPRVLISRPAPRTFAS 240
 DB 234 SADFQIHCAGAKKSGSHAPVRRLDLDALSPQIVLTLEAPRVLISRPAPRTFAS 293
 QY 241 MNMSITKLADKELVHMLISMAKIPGEVELSLFDVRLLESCMMEVIMMLIMKSIHPCK 300
 DB 264 MNMSITKLADKELVHMLISMAKIPGEVELSLFDVRLLESCMMEVIMMLIMKSIHPCK 353
 QY 301 LIFAPDVLVDDEKCKVHGILIEFDMLATTISPRRELKIQHKRYLVKAMILLNSMYPL 360
 DB 354 LIFAPDVLVDDEKCKVHGILIEFDMLATTISPRRELKIQHKRYLVKAMILLNSMYPL 413

CC negatively and for testing environmental pollutants for the presence
 CC of estrogen mimics that may pose health risks involving ER-beta-
 CC mediated processes. The present sequence is human ER-beta protein.

XX Sequence 530 AA:

Query Match 100.0%; Score 2522; Pos 22; Length 530.

Post Local Similarity 100.0%; Prod. No. 2.5e-231; Mismatches 0; Indels 0; Gaps 0;

Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MNSIPSNVNLKGGKQKQTTSPNVLPTPHSHI SPTVHROSHI YAPPOKSPWCAKSL 60
 DB 54 MNSIPSNVNLKGGKQKQTTSPNVLPTPHSHI SPTVHROSHI YAPPOKSPWCAKSL 113
 YY 61 EHTLPVNRKELKKKVSNGKASVPTGSKRPAHPFVAVSYASVHYGVWSSGCKAF 120
 DB 114 EHTLPVNRKELKKKVSNGKASVPTGSKRPAHPFVAVSYASVHYGVWSSGCKAF 173
 YY 121 KRSTGCHREYICPATNOCITDKNPKSCQACRI PKCYEVWVWGSSEFECEVPEVHAF 180
 DB 174 KRSTGCHREYICPATNOCITDKNPKSCQACRI PKCYEVWVWGSSEFECEVPEVHAF 243
 YY 181 SANPQIHACAKKSGSHAPRPVETITDAI SPQIVETLLLEAPRVHLSIPSAFTTAS 240
 DB 234 SANPQIHACAKKSGSHAPRPVETITDAI SPQIVETLLLEAPRVHLSIPSAFTTAS 293
 YY 241 MMSLTKLADKELVHMSMAKKIPGVEVLSLFDQVHLSFCWMEVIMMILMSRTDPRK 300
 DB 294 MMSLTKLADKELVHMSMAKKIPGVEVLSLFDQVHLSFCWMEVIMMILMSRTDPRK 353
 YY 301 LIFAPDVLDRDEKCEVEGLEIFDMILATTSFRELKQHKRYLQVKAMILLNSMPTL 400
 DB 354 LIFAPDVLDRDEKCEVEGLEIFDMILATTSFRELKQHKRYLQVKAMILLNSMPTL 413
 YY 361 VTATQDSSSKRLAHLLNAVTDAI VWTAKSGISSQGSNRIANLMLSHVHANSKRM 420
 DB 414 VTATQDSSSKRLAHLLNAVTDAI VWTAKSGISSQGSNRIANLMLSHVHANSKRM 473
 YY 421 EHTLMKCKNVVYVYDLEMLNAHVLKCKSSITGSESSPALESKAKESGNPQSG 477
 DB 474 EHTLMKCKNVVYVYDLEMLNAHVLKCKSSITGSESSPALESKAKESGNPQSG 530

RESULT 6

AA027322 standard; Protein: 530 AA.

XX AA027322:

XX 18-DEC-2001 (first entry)

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

XX Human (estrogen receptor beta).

DB WPI, 2001 582341765.
 DB R-PSDB: AAS41105; AAS41104.
 PT estrogen receptor gene and protein polymorphisms useful for diagnosis
 PT of individuals at risk of developing bone disorders
 XX Locus: Figure 4; 245pp; English.

The invention relates to a novel isolated peptide comprising or
 CC consisting of an amino acid sequence selected from an amino acid sequence
 CC of a variant estrogen receptor protein (e.g., Ebeta), or a fragment of
 CC 10 amino acids), antibodies against them, nucleic acids encoding
 CC them (including vectors for transforming cells). The gene for human
 CC Ebeta is located on chromosome 17p21. The variant peptides and
 CC proteins can be used in assays to determine the biological
 CC activity of the protein, to raise antibodies, as a reagent in assays
 CC designed to quantitatively determine levels of the protein in
 CC biological fluids, to identify compounds that modulate receptor
 CC activity and to screen compounds for the ability to stimulate or
 CC inhibit interaction between the receptor protein and a target molecule
 CC that normally interacts with the receptor protein (e.g., estrogen).
 CC The antibody can be used to isolate the protein, to assess expression in
 CC disease states e.g., cardiovascular disease and autoimmune disease (e.g.,
 CC systemic lupus erythematosus, arthritis, rheumatism and osteoporosis),
 CC osteoporosis, breast cancer and endometrial cancer. In addition
 CC the antibodies can be used in pharmacogenomic analysis and inhibiting
 CC protein function, e.g., blocking the binding of the estrogen receptor
 CC protein to a binding partner such as a ligand. The nucleic acids
 CC encoding the proteins can be used as probes, primers, chemical
 CC intermediates and in biological assays. The present sequence represents
 CC human Ebeta.

Sequence 530 AA:

Query Match 100.0%; Score 2522; Pos 22; Length 530.

Post Local Similarity 100.0%; Prod. No. 2.5e-231; Mismatches 0; Indels 0; Gaps 0;

Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MNSIPSNVNLKGGKQKQTTSPNVLPTPHSHI SPTVHROSHI YAPPOKSPWCAKSL 60
 DB 54 MNSIPSNVNLKGGKQKQTTSPNVLPTPHSHI SPTVHROSHI YAPPOKSPWCAKSL 113
 YY 61 EHTLPVNRKELKKKVSNGKASVPTGSKRPAHPFVAVSYASVHYGVWSSGCKAF 120
 DB 114 EHTLPVNRKELKKKVSNGKASVPTGSKRPAHPFVAVSYASVHYGVWSSGCKAF 173
 YY 121 KRSTGCHREYICPATNOCITDKNPKSCQACRI PKCYEVWVWGSSEFECEVPEVHAF 180
 DB 174 KRSTGCHREYICPATNOCITDKNPKSCQACRI PKCYEVWVWGSSEFECEVPEVHAF 243
 YY 181 SANPQIHACAKKSGSHAPRPVETITDAI SPQIVETLLLEAPRVHLSIPSAFTTAS 240
 DB 234 SANPQIHACAKKSGSHAPRPVETITDAI SPQIVETLLLEAPRVHLSIPSAFTTAS 293
 YY 241 MMSLTKLADKELVHMSMAKKIPGVEVLSLFDQVHLSFCWMEVIMMILMSRTDPRK 300
 DB 294 MMSLTKLADKELVHMSMAKKIPGVEVLSLFDQVHLSFCWMEVIMMILMSRTDPRK 353
 YY 301 LIFAPDVLDRDEKCEVEGLEIFDMILATTSFRELKQHKRYLQVKAMILLNSMPTL 400
 DB 354 LIFAPDVLDRDEKCEVEGLEIFDMILATTSFRELKQHKRYLQVKAMILLNSMPTL 413
 YY 361 VTATQDSSSKRLAHLLNAVTDAI VWTAKSGISSQGSNRIANLMLSHVHANSKRM 420
 DB 414 VTATQDSSSKRLAHLLNAVTDAI VWTAKSGISSQGSNRIANLMLSHVHANSKRM 473
 YY 421 EHTLMKCKNVVYVYDLEMLNAHVLKCKSSITGSESSPALESKAKESGNPQSG 477
 DB 474 EHTLMKCKNVVYVYDLEMLNAHVLKCKSSITGSESSPALESKAKESGNPQSG 530

RESULT 7

screening of compounds to identify new ligands or hormone analogs which are able to activate the new ER. Chimeric receptors comprising a DBD of the new ER, and LBD and an N-terminal domain derived from another nuclear receptor, can be used to identify new ligands or hormone analogs for the nuclear receptors.

Sequence 540 AA:

Query Match 100.0%; Score 2522; DB 20; Length 540;
Host Local Similarity 100.0%; Pred. No. 2,56-231;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

1 MNSSTSNVNTLEGGGKQRTSPNVLWPTPGHLSPLVHPQLSHLYAEKSPWEASL 60
DB MNSSTSNVNTLEGGGKQRTSPNVLWPTPGHLSPLVHPQLSHLYAEKSPWEASL 113
61 FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 120
DB FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 174
114 FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 174
121 KPSIQGHNDYICPATNGCTTIDNBRKSGYACPLPGCYGVWYGVSTSPRGGYLVPRGR 180
DB KPSIQGHNDYICPATNGCTTIDNBRKSGYACPLPGCYGVWYGVSTSPRGGYLVPRGR 234
174 KPSIQGHNDYICPATNGCTTIDNBRKSGYACPLPGCYGVWYGVSTSPRGGYLVPRGR 234
181 SAHQHLCAGKAKRSGHAPVRELLDALSPEDVITLLEAPRHVLSRISAPFTAS 240
DB SAHQHLCAGKAKRSGHAPVRELLDALSPEDVITLLEAPRHVLSRISAPFTAS 293
241 MMSSTLTKADKELVHMSAKKIPGFVELSLFDQVLLSCMEVLMGLMWRSDIHRGK 300
DB MMSSTLTKADKELVHMSAKKIPGFVELSLFDQVLLSCMEVLMGLMWRSDIHRGK 353
294 MMSSTLTKADKELVHMSAKKIPGFVELSLFDQVLLSCMEVLMGLMWRSDIHRGK 353
301 LIFAPDLVLDROGKCEGILEFPMMLATTSRFRFLKIQHREYLCVAMILLNSMYPL 360
DB LIFAPDLVLDROGKCEGILEFPMMLATTSRFRFLKIQHREYLCVAMILLNSMYPL 413
354 LIFAPDLVLDROGKCEGILEFPMMLATTSRFRFLKIQHREYLCVAMILLNSMYPL 413
361 VTATQADSSRKLAHLINAVTDAVWVIAKSGISSQOOSMRANLMLSHVRAHNSKGM 420
DB VTATQADSSRKLAHLINAVTDAVWVIAKSGISSQOOSMRANLMLSHVRAHNSKGM 473
414 VTATQADSSRKLAHLINAVTDAVWVIAKSGISSQOOSMRANLMLSHVRAHNSKGM 473
421 EHLNKKCKNVVYVYLLLEMLNANVLKCGKSGTSGSPAEDESKSRGSGNPOSG 477
DB EHLNKKCKNVVYVYLLLEMLNANVLKCGKSGTSGSPAEDESKSRGSGNPOSG 540
RESULT 9
ID AAY07270
AAV07270 standard; Protein: 548 AA.
XX
AC AAY07270;
XX
DI 06-JUL-1999 (first entry)
XX
DE Human oestrogen receptor.
XX
KW Human; oestrogen receptor; ligand; bone resorption; metabolic disorder; cardiovascular disease; cancer; central nervous system; breast; uterine; osteoporosis; ovarian; prostate; diabetes; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN W0912961-A1.
XX
PD 18-MAR-1999.
XX
PE 04-SEP-1998; 98WU-0518577.
XX
PR 20-MAR-1998; 98GB-0006032.
XX
PR 08-SEP-1997; 970S-0058271.
XX
PR 30-SEP-1997; 970S-0060520.
XX
PR 30-OCT-1997; 97GB-0022884.
XX
PA (MERL) MERCK & CO INC.

XX
PI Wilkinson B;
XX
PR WPI: 1999-22922/19.
XX
PR N-PSDB: AAX14309.
XX
PT Estrogen receptor
XX
PS Claim 1; Fig 1; 32P; English.

This sequence represents a human oestrogen receptor. The receptor can be used to identify ligands that bind to human oestrogen receptor. The ligands can be used in a method for preventing or treating an oestrogen receptor mediated disease or condition, such as abnormal bone resorption, a cardiovascular disease, cancer, metabolic disorders, or central nervous system disorders. The ligand is especially used to treat osteoporosis, breast, uterine, ovarian or prostate cancer, diabetes or Alzheimer's disease.

Sequence 548 AA:

Query Match 100.0%; Score 2522; DB 20; Length 548;
Host Local Similarity 100.0%; Pred. No. 2,56-231;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

1 MNSSTSNVNTLEGGGKQRTSPNVLWPTPGHLSPLVHPQLSHLYAEKSPWEASL 60
DB MNSSTSNVNTLEGGGKQRTSPNVLWPTPGHLSPLVHPQLSHLYAEKSPWEASL 141
61 FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 120
DB FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 191
132 FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 191
121 KPSIQGHNDYICPATNGCTTIDNBRKSGYACPLPGCYGVWYGVSTSPRGGYLVPRGR 180
DB KPSIQGHNDYICPATNGCTTIDNBRKSGYACPLPGCYGVWYGVSTSPRGGYLVPRGR 251
192 KPSIQGHNDYICPATNGCTTIDNBRKSGYACPLPGCYGVWYGVSTSPRGGYLVPRGR 251
181 SAHQHLCAGKAKRSGHAPVRELLDALSPEDVITLLEAPRHVLSRISAPFTAS 240
DB SAHQHLCAGKAKRSGHAPVRELLDALSPEDVITLLEAPRHVLSRISAPFTAS 311
252 SAHQHLCAGKAKRSGHAPVRELLDALSPEDVITLLEAPRHVLSRISAPFTAS 311
312 MMSSTLTKADKELVHMSAKKIPGFVELSLFDQVLLSCMEVLMGLMWRSDIHRGK 371
301 LIFAPDLVLDROGKCEGILEFPMMLATTSRFRFLKIQHREYLCVAMILLNSMYPL 360
DB LIFAPDLVLDROGKCEGILEFPMMLATTSRFRFLKIQHREYLCVAMILLNSMYPL 411
372 LIFAPDLVLDROGKCEGILEFPMMLATTSRFRFLKIQHREYLCVAMILLNSMYPL 411
361 VTATQADSSRKLAHLINAVTDAVWVIAKSGISSQOOSMRANLMLSHVRAHNSKGM 420
DB VTATQADSSRKLAHLINAVTDAVWVIAKSGISSQOOSMRANLMLSHVRAHNSKGM 491
432 VTATQADSSRKLAHLINAVTDAVWVIAKSGISSQOOSMRANLMLSHVRAHNSKGM 491
421 EHLNKKCKNVVYVYLLLEMLNANVLKCGKSGTSGSPAEDESKSRGSGNPOSG 477
DB EHLNKKCKNVVYVYLLLEMLNANVLKCGKSGTSGSPAEDESKSRGSGNPOSG 548
402 FHITPVNPETI KKVSGNPGASPVTPGSKPRAHPCAVSYASGVYGVWSTCEYKAF 120
RESULT 10
ID ABB80756
ABB80756 standard; Protein: 794 AA.
XX
AC ABB80756;
XX
DI 23-SEP-2002 (first entry)
XX
DE Amino acid sequence of a fusion GFP/beta protein.
XX
KW Fluorometric analyzer; laser; fluorescence; GFP; ER beta; human; gene;
XX
KW green fluorescent protein; estrogen receptor beta; fusion protein.
XX
OS Synthetic.
XX
PA Argentina Victoria.


```

Db      9  MNVSTPSNVINLEKQEGPQTTSPNVLMPITGHLSPVLVHQRSLSHLYAEHPQKSPWCEANSL 68
QY      61  EHTLVNRPETLKRVSNPQASPVTPGSGKPAHFAVGSQDVASCYHWQVMSCEQYKAF 120
        69  EHTLVNRPETLKRVSNPQASPVTPGSGKPAHFAVGSQDVASCYHWQVMSCEQYKAF 128
QY      121  KRSLGCHNDYICPATNCTIDKNRPKSCQACPLKCYEVMWKCSPRECGYPLVRQR 180
        129  KRSLGCHNDYICPATNCTIDKNRPKSCQACPLKCYEVMWKCSPRECGYPLVRQR 188
QY      181  SAEQHLGACAKKAKSGSHAPRVRELLIDALSPQVLTLLEAPPVILISPSAPPTAS 240
        189  SAEQHLGACAKKAKSGSHAPRVRELLIDALSPQVLTLTLLEAPPVILISPSAPPTAS 248
QY      241  MMSLSTKLADKELVHMSWAKKIPGVELSLFGQVHLFSQVMEVIMMGLMMSIDHQR 300
        249  MMSLSTKLADKELVHMSWAKKIPGVELSLFGQVHLFSQVMEVIMMGLMMSIDHQR 308
Db      301  LIFAPDLVLDRDECKVEGILLFIDMLATTSFPRELKQHEEYLCVKAMILLNSSMPL 360
        309  LIFAPDLVLDRDECKVEGILLFIDMLATTSFPRELKQHEEYLCVKAMILLNSSMPL 368
QY      361  VTATQADSSSKLAHLINAVTDALVWYIAKSGISSQGSMLANILMLSHVHASNKM 420
        369  VTATQADSSSKLAHLINAVTDALVWYIAKSGISSQGSMLANILMLSHVHASNKM 428
QY      421  EHLINMKCKNVVPPYDLLEMLNAHVLPCKSSITGSESPADESKSKESQNSQ 477
        429  EHLINMKCKNVVPPYDLLEMLNAHVLPCKSSITGSESPADESKSKESQNSQ 485

```

RESULT 12

AAV01597
ID AAV01597 standard; protein: 485 AA.

XX AAV01597:

XX 17-JUN-1999 (first entry)

XX An oestrogen receptor-beta protein.

XX differential oestrogen receptor: ER-alpha; ER-beta; activation;

KW Apl site; therapy.

XX Homo sapiens.

XX W09911760-A1

XX 11-MAR-1999.

XX 31-AUG-1998: 98W01-0515930.

XX 04-SEP-1997: 97US-0923708.

XX (REGC) UNIV CALIFORNIA

XX Gustafsson J, Kuiper GJM, Kushner PJ, Nilsson S;

XX Paech K, Scanlan TS, Webb P;

XX WPI: 1999-205173/17.

XX N-PSDB: AAX2686.

XX Screening for test compounds which inhibit or activate oestrogen

XX receptor beta (ER-beta) mediated activation at a Apl site - for use

XX in therapy

XX Claim 14; page 43-44; 61pp; English.

XX The present sequence represents an oestrogen receptor-beta. The

XX specification describes a method for screening test compounds for

XX differential oestrogen receptor (ER) alpha-mediated and ER beta-mediated

XX activation at an Apl site. The method is used for identifying compounds

XX Sequence 485 AA:
SO Query Match
98-98: Score 2494; DB 20; Length 485;
Ident. Loc.1 Similarity 99.28; Pred. No. 1,30-228;
Matches 473; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY      1  MRSYTSNVINLEKQEGPQTTSPNVLMPITGHLSPVLVHQRSLSHLYAEHPQKSPWCEANSL 60
        68  EHTLVNRPETLKRVSNPQASPVTPGSGKPAHFAVGSQDVASCYHWQVMSCEQYKAF 120
        9  MNVSTPSNVINLEKQEGPQTTSPNVLMPITGHLSPVLVHQRSLSHLYAEHPQKSPWCEANSL 68
Db      61  EHTLVNRPETLKRVSNPQASPVTPGSGKPAHFAVGSQDVASCYHWQVMSCEQYKAF 120
        69  EHTLVNRPETLKRVSNPQASPVTPGSGKPAHFAVGSQDVASCYHWQVMSCEQYKAF 128
QY      121  KRSLGCHNDYICPATNCTIDKNRPKSCQACPLKCYEVMWKCSPRECGYPLVRQR 180
        129  KRSLGCHNDYICPATNCTIDKNRPKSCQACPLKCYEVMWKCSPRECGYPLVRQR 188
Db      181  SAEQHLGACAKKAKSGSHAPRVRELLIDALSPQVLTLLEAPPVILISPSAPPTAS 240
        189  SAEQHLGACAKKAKSGSHAPRVRELLIDALSPQVLTLTLLEAPPVILISPSAPPTAS 248
QY      241  MMSLSTKLADKELVHMSWAKKIPGVELSLFGQVHLFSQVMEVIMMGLMMSIDHQR 300
        249  MMSLSTKLADKELVHMSWAKKIPGVELSLFGQVHLFSQVMEVIMMGLMMSIDHQR 308
Db      301  LIFAPDLVLDRDECKVEGILLFIDMLATTSFPRELKQHEEYLCVKAMILLNSSMPL 360
        309  LIFAPDLVLDRDECKVEGILLFIDMLATTSFPRELKQHEEYLCVKAMILLNSSMPL 368
QY      361  VTATQADSSSKLAHLINAVTDALVWYIAKSGISSQGSMLANILMLSHVHASNKM 420
        369  VTATQADSSSKLAHLINAVTDALVWYIAKSGISSQGSMLANILMLSHVHASNKM 428
QY      421  EHLINMKCKNVVPPYDLLEMLNAHVLPCKSSITGSESPADESKSKESQNSQ 477
        429  EHLINMKCKNVVPPYDLLEMLNAHVLPCKSSITGSESPADESKSKESQNSQ 485

```

RESULT 13

AAW98125
ID AAW98125 standard; protein: 549 AA.

XX AAW98125:

XX 21-JUN-1999 (first entry)

XX Mouse oestrogen receptor beta isoform mER beta-1.

XX oestrogen receptor, ER beta 1, ER beta c, mouse; tumour;

XX diagnosis; therapy; isoform.

XX Mus musculus.

XX W09905170-A1.

XX 04-FEB-1999.

XX 28-JUL-1998: 98W01-0515539.

XX 30-JUL-1997: 97US-0054210.

XX 28-JUL-1997: 97US-0053869.

XX (UYA) UNIV YALE.

XX Denton R, Fraslon C, Galien R, Rischio-Ridon M.

XX WPI: 1999-14282/12.

XX Human oestrogen receptor beta-3 - useful in treating and monitoring

XX tumours

Example 2: Fig 2a: 89pp: English.

XX The present sequence is alternatively spliced isoform mkr beta-1
 CC of the murine oestrogen receptor-beta (ER beta-c) or
 CC mkr beta-3). The mkr beta-4 gene (see AAX24945) contains 9 exons,
 CC including the newly described exon 5b, which is not found in ER
 CC beta-1. Exon 5b is located in the ligand binding domain of mkr
 CC beta-4, and probably plays a significant role in mkr beta-4 function.
 CC The mkr beta-1 clone was isolated by RT-PCR of mRNA, and by
 CC amplification from a mouse embryonic stem cell genomic DNA library.
 CC It lacks exon 5b. Preliminary data indicate that the mkr beta-1
 CC isoform may be more active than the full-length mkr beta-4 (see
 CC AAW98126). 2 other isoforms, mkr beta-2 (see AAW98126) and rat ER
 CC beta-4 (see AAW98127), have also been isolated, and the human exon 5b
 CC region (see AAX24944) has been identified. Based on these sequences,
 CC the invention provides methods for identifying agents that block or
 CC augment ER beta-c mediated transcriptional regulation, methods to
 CC determine whether ER beta-c or its isoforms is being expressed in
 CC tissues or cells, and methods of identifying and using agents that
 CC block the transcriptional regulation of genes by ER beta-c or its
 CC isoforms, which in turn modulates of other biological and pathological
 CC processes. Gene therapy expression of ER beta-c is envisaged.

XX Sequence: 549 AA:

Query Match: 89.78% Score: 2257; DB: 20; Length: 549;
 Best Local Similarity: 88.78% Pred. No. 4.7e-206;

Matches: 424; Conservative: 22; Mismatches: 41; Indels: 0; Gaps: 0.

1 MNSVSPSNVNI ECTGPGQTSPNVI WPTGSHSPI VVHPSHIVAFPKSPWFAVSI 60
 73 MNSVSPSTGNIECTGPGQTSPNVI WPTGSHSPI VVHPSHIVAFPKSPWFAVSI 132
 61 EHTLPVNRKELKKKVSNGKASVYTGSKRDHAFVAVSDVASGYHGVMSSTEGKAF 120
 133 EHTLPVNRKELKKKVSNGKASVYTGSKRDHAFVAVSDVASGYHGVMSSTEGKAF 192
 121 KRSLGSHNDYCTATNGCTTDKNNPKSCVACPLKCYEVGVWVCGSPRPGCYPLVPPGP 180
 193 KRSLGSHNDYCTATNGCTTDKNNPKSCVACPLKCYEVGVWVCGSPRPGCYPLVPPGP 252
 181 SADBOLHAGAKAKSGSHAPRVELLDALSPGVALTLLEAPPPVLLSRPSAPPTAS 240
 253 SASDVGELGAKAKSGSHAPRVELLDALSPGVALTLLEAPPPVLLSRPSAPPTAS 312
 241 MMSSTLTKLAKELVHMI SNAKKIPGVEVLSLPGVYLLSSWMEVIMMLIMKRSIDHCK 300
 313 MMSSTLTKLAKELVHMI SNAKKIPGVEVLSLPGVYLLSSWMEVIMMLIMKRSIDHCK 372
 401 LLEAPDLVLDRECKVEGILLETEDMLATTSREBELKIDHKEYLGVKAMILLNSMPTL 460
 473 LLEAPDLVLDRECKVEGILLETEDMLATTSREBELKIDHKEYLGVKAMILLNSMPTL 432
 461 VTATGDDSSRKLAHLINAVTDAI VVVI AKSGTSSQGSQSMRLANILMLSHVRIASNKGM 420
 433 ATASQFAESSRKLTTHILINAVTDAI VVVI AKSGTSSQGSQSMRLANILMLSHVRIASNKGM 492
 421 EHTLNMCKKNVVVYVYLLLEMLNANHVIKPKKSTTSSESPADNSKSGSNGPQSO 477
 493 EHTLNMCKKNVVVYVYLLLEMLNANHVIKPKKSTTSSESPADNSKSGSNGPQSO 549

RESULT 14

AAV04444 standard: protein: 549 AA.

AAV04444.

25 JUN 1999 (first entry)

Murine mkr beta-1 clone protein sequence.

Estrogen receptor beta; oestrogen receptor beta; ER beta.

XX Mus sp.
 CC W09905171-A1.
 CC 04-FeB-1999.
 CC 28-JUL-1998; 98W0-0815540.
 CC 40-JUL-1997; 97US-0054210.
 CC 28-JUL-1997; 97US-0053869.
 CC (UYA) UNIV YALE.
 CC Baron R, Chambon P, Denton R, Dupont S, Garrier J;
 CC W01: 1999-142853/12.

Novel murine oestrogen receptor-beta genes and splice variants -
 useful for treating oestrogen receptor-beta mediated disorders

Claim 9: Fig 2a: 89pp: English.

XX The present invention describes the murine complete oestrogen receptor
 CC beta (mkr beta-c) also designated mkr beta-3). The present sequence
 CC represents mkr beta-1, which is an isoform of mkr beta-c. The protein
 CC can be used to screen for agents that modulate or block the interaction
 CC between the ER beta-c protein and physiological, allhene or
 CC triphenylethylene (ant) oestrogens. The antibody which binds the protein
 CC can be used in assays to detect activation of the ER beta protein and
 CC can be used for subcellular localisation of the protein. The
 CC complementary nucleic acid can be used to inhibit expression of the
 CC ER beta c nucleic acid.

XX Sequence: 549 AA:

Query Match: 89.78% Score: 2255; DB: 20; Length: 549;
 Best Local Similarity: 88.78% Pred. No. 7.4e-206;

Matches: 424; Conservative: 22; Mismatches: 41; Indels: 0; Gaps: 0;

1 MNSVSPSNVNI ECTGPGQTSPNVI WPTGSHSPI VVHPSHIVAFPKSPWFAVSI 60
 73 MNSVSPSTGNIECTGPGQTSPNVI WPTGSHSPI VVHPSHIVAFPKSPWFAVSI 132
 61 EHTLPVNRKELKKKVSNGKASVYTGSKRDHAFVAVSDVASGYHGVMSSTEGKAF 120
 133 EHTLPVNRKELKKKVSNGKASVYTGSKRDHAFVAVSDVASGYHGVMSSTEGKAF 192
 121 KRSLGSHNDYCTATNGCTTDKNNPKSCVACPLKCYEVGVWVCGSPRPGCYPLVPPGP 180
 193 KRSLGSHNDYCTATNGCTTDKNNPKSCVACPLKCYEVGVWVCGSPRPGCYPLVPPGP 252
 181 SADBOLHAGAKAKSGSHAPRVELLDALSPGVALTLLEAPPPVLLSRPSAPPTAS 240
 253 SASDVGELGAKAKSGSHAPRVELLDALSPGVALTLLEAPPPVLLSRPSAPPTAS 312
 241 MMSSTLTKLAKELVHMI SNAKKIPGVEVLSLPGVYLLSSWMEVIMMLIMKRSIDHCK 300
 313 MMSSTLTKLAKELVHMI SNAKKIPGVEVLSLPGVYLLSSWMEVIMMLIMKRSIDHCK 372
 401 LLEAPDLVLDRECKVEGILLETEDMLATTSREBELKIDHKEYLGVKAMILLNSMPTL 460
 473 LLEAPDLVLDRECKVEGILLETEDMLATTSREBELKIDHKEYLGVKAMILLNSMPTL 432
 461 VTATGDDSSRKLAHLINAVTDAI VVVI AKSGTSSQGSQSMRLANILMLSHVRIASNKGM 420
 433 ATASQFAESSRKLTTHILINAVTDAI VVVI AKSGTSSQGSQSMRLANILMLSHVRIASNKGM 492
 421 EHTLNMCKKNVVVYVYLLLEMLNANHVIKPKKSTTSSESPADNSKSGSNGPQSO 477
 493 EHTLNMCKKNVVVYVYLLLEMLNANHVIKPKKSTTSSESPADNSKSGSNGPQSO 549

RESULT 15
AAW14723
ID AAW14723 standard; Protein: 485 AA.
XX
XX AAW14723:
XX
XX 08-JUN-1997 (first entry)
XX
XX Rat oestrogen receptor beta (ER-beta).
XX
XX orphan receptor; oestrogen receptor beta; ER-beta;
XX nuclear receptor; prostate cancer; benign prostatic hyperplasia;
XX ovary cancer; cardiovascular disease; osteoporosis;
XX environment; pollutant.
XX
XX Rattus sp.
XX
XX Location: % alignment
XX 104..169
XX Domain
XX 259..457
XX 21361.. Ligand binding domain
XX
XX W09709448-A2.
XX
XX 14-MAR-1997.
XX
XX 09-SEP-1996; 96WD-EP03943.
XX
XX 08-MAY-1996; 96GB-0009576.
XX 08-SEP-1995; 95GB-0018272.
XX 15-MAR-1996; 96GB-0005550.
XX 11-APR-1996; 96GB-0007542.
XX
XX (KARO-) KAKO B10 AB.
XX
XX Emmark E, Gustafsson J, Kuiper GG;
XX
XX WP1: 1997-192842/17.
XX
XX N-PSDB; AAT62842.
XX
XX New isolated oestrogen receptor beta - used to develop prods. for
XX treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
XX disease.
XX
XX Claim 1: Fig 1: 45pp; English.
XX
XX A novel rat orphan receptor (AAW14723) is related to the known
XX oestrogen receptor ER-alpha, and has been designated ER-beta.
XX It is an oestrogen receptor-related nuclear receptor. The ER-beta
XX amino acid sequence was deduced from a cDNA clone (AAT62842) isolated
XX from a rat prostate cDNA library. Rat, human and mouse ER-beta
XX (AAW14723-25) can be used to isolate molecules for use in the
XX treatment of cardiovascular diseases, central nervous system
XX diseases, osteoporosis, prostate or ovarian cancer or benign
XX prostatic hyperplasia and to test environmental chemicals for
XX oestrogenic activity.
XX
XX Sequence 485 AA:
XX
XX Query Match: 89.2%, Score 2249; DB 18; Length 485;
XX Best local similarity: 88.5%, Freq. No. 2,36 205;
XX Matches 422; Constructive 24; Mismatches 32; Indels 0; Gaps 0;
XX
XX 1 MNYSTFSSVNTI FCGPRGRTSPNVLPPTAH SPIVHQLSHEVAEPSEFWPARSL 60
XX IIIII : III : III : IIIIIIII IIIII : : IIIIIIII IIIII
XX 9 MNYSPGSTSNIDGGVRLSLSPVIMPTSGHLSPLATH QSSLLAEPKSKWCARSL 68
XX
XX 61 EHTLPVNRFTI KRRVSNRPAAPVTGSGKPDHHPAVSIVASGYHYGVWSCGGKAPF 120
XX IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 69 EHTLPVNRFTI KPRISGVASAPVSPAKPDHHPVTSIVASGYHYGVWSCGGKAPF 128
XX
XX 121 KSLIGHNIVLPATINCTIHKNPGRSGQAPRI PRVEVAVWVGSGRHRGCVIRVRQR 180

RESULT 16
AAW01596
ID AAW01596 standard; Protein: 485 AA.
XX
XX AAW01596:
XX
XX 17-JUN-1999 (first entry)
XX
XX An estrogen receptor-beta protein.
XX
XX differential estrogen receptor; ER-alpha; ER-beta; activation;
XX API site; therapy.
XX
XX Rattus sp.
XX
XX W09911760-A1.
XX
XX 11-MAR-1999.
XX
XX 31-AUG-1998; 98WD-US18030.
XX
XX 04-SEP-1997; 97US-0923708.
XX
XX (FESC-) UNIV CALIFORNIA.
XX
XX Gustafsson J, Kuiper GJM, Kushner PJ, Nilsson S;
XX Paech K, Scanlan TS, Webb P;
XX
XX WP1: 1999-206173/17.
XX
XX R-TRDB; AAX26685.
XX
XX Screening for test compounds which inhibit or activate estrogen
XX receptor beta (ER beta) mediated activation at a API site for use
XX in therapy.
XX
XX Claim 3: Page 41-43; 61pp; English.
XX
XX The present sequence represents an estrogen receptor beta. The
XX specification describes a method for screening test compounds for
XX differential estrogen receptor (ER) alpha mediated and ER beta mediated
XX activation at an API site. The method is used for identifying compounds
XX which can be used in therapy.
XX
XX Sequence 485 AA:
XX
XX Query Match: 89.2%, Score 2249; DB 20; Length 485;
XX Best local similarity: 88.5%, Freq. No. 2,36 205;
XX Matches 422; Constructive 24; Mismatches 32; Indels 0; Gaps 0;
XX
XX 129 KSLIGHNIVLPATINCTIHKNPGRSGQAPRI PRVEVAVWVGSGRHRGCVIRVRQR 188
XX IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 181 SAEQLHPCAGAPRSCSHAPPEVELLFAISPEQLVLIILFAEFGVLSPEAPETPAS 240
XX I : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 189 SSSPOVHCLSKAKNKGSHAPRVEKLLSTSPEDIVITLI EAPDPVAVISRSMPTTAS 248
XX
XX 241 MMSSTIPLADRELVHMLSMARIPGVESLPIVPLLSGQWFEVIMLMPSTPDDCK 300
XX IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 249 MMSSTIPLADRELVHMLSMARIPGVESLPIVPLLSGQWFEVIMLMPSTPDDCK 308
XX
XX 301 ELPFQVILGGESEKVEEILLETIOMLATISPEPEKIGREFYGVKAMIIINSMPPL 360
XX IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 309 IIFADPLVLPDESKVEGILLETIOMLATISPEPEKIGREFYGVKAMIIINSMPPL 368
XX
XX 361 VTATQAESSPKELMLNAVILAVVAVIARSGISSQSGMGLANMLLSHVHASNKM 420
XX I : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 363 ASANAFASRRI LQD NAVILAVVAVIARSGISSQSGVRI ANIMLSHVHINRKM 428
XX
XX 421 EHTLPVNRFTI KRRVSNRPAAPVTGSGKPDHHPAVSIVASGYHYGVWSCGGKAPF 477
XX IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII : IIIII
XX 423 EHTLPVNRFTI KPRISGVASAPVSPAKPDHHPVTSIVASGYHYGVWSCGGKAPF 485

XX Claim 4: Page 20-22; 45pp; English.

ES This partial sequence represents a splice variant of a novel estrogen
XX binding protein, detected by screening a human testis cDNA library.

CC This protein contains an alternative exon 8 (exon 8B) of the novel
CC estrogen receptor represented. As a consequence of the introduction of
CC this exon through an alternative splicing reaction, the reading frame
CC encoding the novel receptor is immediately terminated, creating a
CC truncation of the carboxy terminus of the novel receptor. This receptor
CC is able to bind and be activated by estradiol, estrone and estrifol, can be
CC used in a screening assay for the identification of new drugs e.g. novel
CC ligands or hormonal analogues. This variant does not contain an AF 2
CC region and therefore probably lacks the ability to modulate transcription
CC or target genes in a ligand dependent fashion.

XX Sequence 416 AA:

Query Match 87.2%; Score 2198; DB 18; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.3e-200;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNVSPSNVNI FQAPGPTTSPNVI WPTPGH SPI VVHQI SH YAFYKSHWFEARSL 60
DB 1 MNVSPSNVNI FQAPGPTTSPNVI WPTPGH SPI VVHQI SH YAFYKSHWFEARSL 60
QY 61 EHTLPVNRETLKRPVSGNPGASVPTGSKRIAHFVAVSGVASYHYGVMSFEQKAF 120
DB 61 EHTLPVNRETLKRPVSGNPGASVPTGSKRIAHFVAVSGVASYHYGVMSFEQKAF 120
QY 121 KPSIQGNHYTCTATNCTIDKNPKRSYACQAFKRYGVYMWYCSFPGCYRFLVVROR 180
DB 121 KPSIQGNHYTCTATNCTIDKNPKRSYACQAFKRYGVYMWYCSFPGCYRFLVVROR 180
QY 181 SAEQIHQAKKAKSGGAPRVRELLDALSPQIVLITLFAHPVILISPSAPTEAS 240
DB 181 SAEQIHQAKKAKSGGAPRVRELLDALSPQIVLITLFAHPVILISPSAPTEAS 240
QY 241 MNVSTIKLAIKELVHMI SMARKIPGVNELSLFQVRLPSQWMEVIMM1MMS110HPK 400
DB 241 MNVSTIKLAIKELVHMI SMARKIPGVNELSLFQVRLPSQWMEVIMM1MMS110HPK 400
QY 301 LIFAPDLVLDDECKVEGIIETFDMLATTSPEPEIKIQRKYGVVKAMILLNSMYP 460
DB 301 LIFAPDLVLDDECKVEGIIETFDMLATTSPEPEIKIQRKYGVVKAMILLNSMYP 460
QY 361 VTATQDADSSKRLAHLNAVTDAI VWTAKSGISSQOQSRANLIMLSHVHA 415
DB 361 VTATQDADSSKRLAHLNAVTDAI VWTAKSGISSQOQSRANLIMLSHVHA 415

RESULT 22
AAH47837
ID AAH47837 standard; Protein: 416 AA.

XX AAH47837:
XX 25-MAR-2002 (first entry)
XX
XX ER splice variant ER-beta.
XX
XX DNA binding domain; DBD; ligand binding domain; LBD; chimeric receptor;
XX estrogen receptor; ER; chimeric; 14 ER-alpha; ER-beta; ER-gamma;
XX estradiol; nuclear receptor; progesterone receptor; amplification; primer;
XX polymerase chain reaction; AF-2 region; splice variant.

OS Homo sapiens.
XX
XX
XX EP1162264-A2.
XX
XX 12-DEC-2001.
XX
XX 25-MAR-1997; 2001EP-0202021.

XX 26-MAR-1996; 96EP-0200820.
XX
XX 22-NOV-1996; 96EP-0203284.
XX 25-MAR-1997; 97EP-0200903.
XX
XX (ALKU) AKZO NOBEL NV.
XX
XX Mosselman S, Dijkema R;
XX
XX WIL: 2002-084414/12;
XX N-PSDB; AA172153.
XX
XX New isolated chimeric receptor comprising a DNA binding domain and/or
XX ligand binding domain of a new estrogen receptor, for identifying
XX functional ligands or hormonal analogs for the receptor.

PS Example B: Page 17-18; 45pp; English.

XX The sequences given in AAH47836 and AAH47837 show splice variants
XX of a novel estrogen receptor (ER). The gene encoding this new ER is
XX located on chromosome 14 and has a different tissue distribution from
XX classical ER. This ER also has two orphan ER's, ER-alpha and ER-beta.
XX These orphan receptors have estrogen receptor related structure but do
XX not appear to be able to bind estradiol or other ER ligands. These
XX splice variants differ from the parent ER around exon 8, one clone
XX contains exon 8B through alternative splicing, thus causes an immediate
XX termination of the reading thereby creating a truncation at the carboxy
XX terminus. The other splice variant contains an alternative exon 8, exon
XX 8C, which encodes two C terminal amino acids and then has a stop codon.
XX The splice variant proteins do not contain the AF 2 region and therefore
XX probably lack the ability to modulate transcription of target genes in
XX a ligand dependent manner. However they may be able to interfere with
XX the functioning of the WT classical ER or the novel ER of the invention,
XX either by heterodimerisation or by occupying estrogen response elements
XX or by interactions with other transcription factors.

XX Sequence 416 AA:

Query Match 87.2%; Score 2198; DB 23; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.3e-200;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNVSPSNVNI FQAPGPTTSPNVI WPTPGH SPI VVHQI SH YAFYKSHWFEARSL 60
DB 1 MNVSPSNVNI FQAPGPTTSPNVI WPTPGH SPI VVHQI SH YAFYKSHWFEARSL 60
QY 61 EHTLPVNRETLKRPVSGNPGASVPTGSKRIAHFVAVSGVASYHYGVMSFEQKAF 120
DB 61 EHTLPVNRETLKRPVSGNPGASVPTGSKRIAHFVAVSGVASYHYGVMSFEQKAF 120
QY 121 KPSIQGNHYTCTATNCTIDKNPKRSYACQAFKRYGVYMWYCSFPGCYRFLVVROR 180
DB 121 KPSIQGNHYTCTATNCTIDKNPKRSYACQAFKRYGVYMWYCSFPGCYRFLVVROR 180
QY 181 SAEQIHQAKKAKSGGAPRVRELLDALSPQIVLITLFAHPVILISPSAPTEAS 240
DB 181 SAEQIHQAKKAKSGGAPRVRELLDALSPQIVLITLFAHPVILISPSAPTEAS 240
QY 241 MNVSTIKLAIKELVHMI SMARKIPGVNELSLFQVRLPSQWMEVIMM1MMS110HPK 400
DB 241 MNVSTIKLAIKELVHMI SMARKIPGVNELSLFQVRLPSQWMEVIMM1MMS110HPK 400
QY 301 LIFAPDLVLDDECKVEGIIETFDMLATTSPEPEIKIQRKYGVVKAMILLNSMYP 460
DB 301 LIFAPDLVLDDECKVEGIIETFDMLATTSPEPEIKIQRKYGVVKAMILLNSMYP 460
QY 361 VTATQDADSSKRLAHLNAVTDAI VWTAKSGISSQOQSRANLIMLSHVHA 415
DB 361 VTATQDADSSKRLAHLNAVTDAI VWTAKSGISSQOQSRANLIMLSHVHA 415

RESULT 24
AAH37214


```

61 EHTLPVNRKELKPKVSNPQASVVTGPKSKQAPASVSTVYASYYWVWSEGYKAF 120
|||||
61 EHTLPVNRKELKPKVSNPQASVVTGPKSKQAPASVSTVYASYYWVWSEGYKAF 120
|||||
121 KRSLGCHNNYTGATWCTTQKPKPSQACPLPKCYEYGVKPKSSPEPPGYRVPKQR 180
|||||
121 KRSLGCHNNYTGATWCTTQKPKPSQACPLPKCYEYGVKPKSSPEPPGYRVPKQR 180
|||||
181 SAEQGHQACAKAPSSSHAPVPELILDAISPEQVITLLEAPPHVILSPISAPPTFAS 240
|||||
181 SAEQGHQACAKAPSSSHAPVPELILDAISPEQVITLLEAPPHVILSPISAPPTFAS 240
|||||
241 MMSSTIKLAKELVHMIISMAKKIPGFVELSLFQVRLLESQWMEVLMGLMWSIDHCK 300
|||||
241 MMSSTIKLAKELVHMIISMAKKIPGFVELSLFQVRLLESQWMEVLMGLMWSIDHCK 300
|||||
301 LIFAPDVLDRPKQVEGILFEDMLATTSPEPKLQHKKEYLCVKAMILLNSSMYPL 460
|||||
301 LIFAPDVLDRPKQVEGILFEDMLATTSPEPKLQHKKEYLCVKAMILLNSSMYPL 460
|||||
361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMKLANILMLSHVRIA 415
|||||
361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMKLANILMLSHVRIA 415
|||||

RESULT 25
AAV04435
10 AAV04435 standard; Protein: 510 AA
AAV04435:
25-JUN-1999 (first entry)
XX
XX Murine mkr beta-2 clone protein sequence.
XX
XX Estrogen receptor beta; estrogen receptor beta; ER beta.
XX
XX Mus sp
XX
XX W09905171-A1
XX
XX 04-FEB-1999.
XX
XX 28-JUL-1998: 98W0-051540.
XX
XX 30-JUL-1997: 970S-0054210.
XX
XX 28-JUL-1997: 970S-0054869.
XX
XX (UYVA ) UNIV YALE.
XX
XX Barton K, Chabon P, Denton R, Dupont S, Garnier J.
XX
XX WPI: 1999-142853/12.
XX
XX Novel murine estrogen receptor beta genes and splice variants -
XX
XX useful for treating estrogen receptor-beta mediated disorders
XX
XX Claim 9; Fig 2b; 89pp; English.
XX
XX The present invention describes the murine complete estrogen receptor
XX
XX beta (mkr beta c) also designated mkr beta-3). The present sequence
XX
XX represents mkr beta-2, which is an isoform of mkr beta c. The protein
XX
XX can be used to screen for agents that modulate or block the interaction
XX
XX between the ER beta c protein and physiological, silicone or
XX
XX triphenylethylene (anti)estrogens. The antibody which binds the protein
XX
XX can be used in assays to detect activation of the ER beta protein and
XX
XX complementary nucleic acid can be used to inhibit expression of the
XX
XX ER beta c nucleic acid.
XX
XX Sequence 510 AA:

```

```

Query Match 74.4% Score 2000.50 Db = 20011
Seqs Used: 1 similarity 60.7% from 1,26-181
Matches: 222 Conservative 222 Mismatches 31 Gaps 1
Gaps 1:

QY 1 MMSSTIKLAKELVHMIISMAKKIPGFVELSLFQVRLLESQWMEVLMGLMWSIDHCK 400
DB 74 MMSSTIKLAKELVHMIISMAKKIPGFVELSLFQVRLLESQWMEVLMGLMWSIDHCK 400
QY 121 KRSLGCHNNYTGATWCTTQKPKPSQACPLPKCYEYGVKPKSSPEPPGYRVPKQR 180
DB 121 KRSLGCHNNYTGATWCTTQKPKPSQACPLPKCYEYGVKPKSSPEPPGYRVPKQR 180
QY 181 SAEQGHQACAKAPSSSHAPVPELILDAISPEQVITLLEAPPHVILSPISAPPTFAS 240
DB 181 SAEQGHQACAKAPSSSHAPVPELILDAISPEQVITLLEAPPHVILSPISAPPTFAS 240
QY 241 MMSSTIKLAKELVHMIISMAKKIPGFVELSLFQVRLLESQWMEVLMGLMWSIDHCK 300
DB 241 MMSSTIKLAKELVHMIISMAKKIPGFVELSLFQVRLLESQWMEVLMGLMWSIDHCK 300
QY 301 LIFAPDVLDRPKQVEGILFEDMLATTSPEPKLQHKKEYLCVKAMILLNSSMYPL 460
DB 301 LIFAPDVLDRPKQVEGILFEDMLATTSPEPKLQHKKEYLCVKAMILLNSSMYPL 460
QY 361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMKLANILMLSHVRIA 415
DB 361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMKLANILMLSHVRIA 415
QY 421 EHTLPVNRKELKPKVSNPQASVVTGPKSKQAPASVSTVYASYYWVWSEGYKAF 120
DB 421 EHTLPVNRKELKPKVSNPQASVVTGPKSKQAPASVSTVYASYYWVWSEGYKAF 120
QY 454 EHTLPVNRKELKPKVSNPQASVVTGPKSKQAPASVSTVYASYYWVWSEGYKAF 120
DB 454 EHTLPVNRKELKPKVSNPQASVVTGPKSKQAPASVSTVYASYYWVWSEGYKAF 120

RESULT 26
AAW98126
10 AAW98126 standard; Protein: 510 AA.
AAW98126:
21-JUN-1999 (first entry)
XX
XX Mouse estrogen receptor beta isoform mkr beta-2.
XX
XX Estrogen receptor; ER beta-2; ER beta c; mouse; tumour;
XX
XX diagnosis; therapy; isoform.
XX
XX Mus musculus.
XX
XX W09905170-A1.
XX
XX 04-FEB-1999.
XX
XX 28-JUL-1998: 98W0-051549.
XX
XX 30-JUL-1997: 970S-0054210.
XX
XX 28-JUL-1997: 970S-0054869.
XX
XX (UYVA ) UNIV YALE.
XX
XX Denton K, Fraston C, Gallen K, Rischke-Rigon M.
XX
XX WPI: 1999-142852/12.
XX
XX Human estrogen receptor beta-3 - useful in treating and monitoring
XX
XX tumours
XX
XX Example 2; Fig 2b; 89pp; English.
XX
XX The present sequence is alternatively spliced isoform mkr beta 2

```

of the murine complete oestrogen receptor-beta (ER beta-c), or
 cc mER beta-4). The mER beta-4 gene (see AAX24945) contains 9 exons,
 cc including the newly described exon 5b, which is not found in ER
 cc beta-1. Exon 5b is located in the 11and binding domain of mER
 cc beta-1, and probably plays a significant role in mER beta-4 function.
 cc The mER beta-2 clone was isolated from an embryonic stem cell
 cc genomic library. It lacks exon 3, 2 other isoforms, mER beta-1
 cc (see AAW98127) and rat ER beta-4 (see AAW98127), have also been
 cc isolated, and the human exon 5b region (see AAX24944) has been
 cc identified. Based on these sequences, the invention provides
 cc methods for identifying agents that block or augment ER beta-c
 cc mediated transcriptional regulation, methods to determine whether
 cc ER beta-c or its isoforms is being expressed in tissues or cells,
 cc and methods of identifying and using agents that block the
 cc transcriptional regulation of genes by ER beta-c or its isoforms,
 cc which in turn modulates other biological and pathological
 cc processes. Gene therapy expression of ER beta-c is envisaged.

cc Sequence 510 AA:

Query Match 79.4%; Score 2000.5; DB 20; Length 510;
 Post Local Similarity 80.7%; Prod No 1 20 181;
 Matches 385; Conservative 22; Mismatches 31; Indels 39; Gaps 1.

QY 1 MNTSIPSNVNLGCGRGKRTSPNVLWPTGHSPLVVRHQLSHLYAEFGSPWCEARSL 60
 DB 73 MNTSVSVSTNLEGGVGRQVSPNVLWPTGHSPLVVRHQLSHLYAEFGSPWCEARSL 132
 QY 61 EHLLVRRKELKKRVSNGPVSPTGSKRPAHPAVNSIVASGYHGVMSCEGCKAF 120
 DB 133 EHLLVRRKELKKRVSNGPVSPTGSKRPAHPAVNSIVASGYHGVMSCEGCKAF 192
 QY 121 KRSTGHNQYTGATNQTGDKNRKRSQVAPLPRVEYGVWVCSPEEPGYPLVRQR 180
 DB 193 KRSTGHNQYTGATNQTGDKNRKRSQVAPLPRVEYGVWVCSPEEPGYPLVRQR 213
 QY 181 SADEQDLGACAKKRSCHAPRVRELLDALSPQVLTLLFAPPHVLTLSRSAPETAS 240
 DB 214 SASQVGGTNRKAKRTSHTPRVRELLDALSPQVLTLLFAPPHVLTLSRSAPETAS 273
 QY 241 MMSSTKLAKRELVIHMSAKKIPGVVELSTFGVYHLEFSEWMEVIMMGLMKRSIDHCK 300
 DB 274 MMSSTKLAKRELVIHMSAKKIPGVVELSTFGVYHLEFSEWMEVIMMGLMKRSIDHCK 333
 QY 401 LIFAPDLVLDKDEKVEGILLLEPMLATTSPEPKLKQIKRYLVKAMILLNSSTYL 460
 DB 434 LIFAPDLVLDKDEKVEGILLLEPMLATTSPEPKLKQIKRYLVKAMILLNSSTYL 493
 QY 461 VTFQDASSPKLAHLNAVTVQVWVAFKSTSSQSGMPLANLMLSEVPHASNECM 420
 DB 494 ATNSQGESESRKLTILLNAVTVQVWVAFKSTSSQSGMPLANLMLSEVPHASNECM 453
 QY 421 EHLLNMKCKNVAVVYDILLPMNAHVLKQKSSITSSPSPAHSKSPKSNPSSQ 477
 DB 454 EHLLNMKCKNVAVVYDILLPMNAHVLKQKSSITSSPSPAHSKSPKSNPSSQ 510

RESULT 27

AAW98127 standard; Protein: 414 AA.

AAW98127:

21 JUN 1999 (first entry)

rat oestrogen receptor beta isoform mER beta-4.

oestrogen receptor: ER beta-4; ER beta-c; rat; tumour;

diagnosis; therapy; isoform.

catious not verified.

W09905170-A1.

XX 04 FEB 1999.
 PD 26 JUL 1998; 9800-0815539.
 FF 30 JUL 1997; 9708-0054210.
 PR 26 JUL 1997; 9708-0053869.
 XX (UYVA) UNIV YALE.
 PA Denton R, Frasion C, Gallon R, Rische-Kron M;
 PI WPL: 1997-1428-2/12.
 PS Example 2: F14 2c; 89pp; English.

XX The present sequence is alternatively spliced isoform ER beta-4
 cc of the complete oestrogen receptor beta (ER beta-c). Murine ER
 cc beta-c clone mER beta-4 (see AAX24945) has been identified, which
 cc includes newly described exon 5b, which is located in the 11and
 cc binding domain of the receptor and is probably involved in receptor
 cc function. ER beta-4 nucleic acid was obtained from rat ovaries.
 cc It includes exon 5b but lacks exon 6, causing a frameshift and
 cc truncation of the protein 13 amino acids beyond the translated
 cc exon 5b. Also disclosed are mER beta-1 (AAW98125) and mER beta-2
 cc (AAW9129) isoforms, and human exon 5b (AAW24944). Based on these
 cc sequences, the invention provides methods for identifying agents
 cc that block or augment ER beta-c mediated transcriptional regulation,
 cc methods to determine whether ER beta-c or its isoforms is being
 cc expressed in tissues or cells, and methods of identifying and using
 cc agents that block the transcriptional regulation of genes by ER
 cc beta-c or its isoforms, which in turn modulates other biological
 cc and pathological processes. Gene therapy expression of ER beta-c
 cc is envisaged.

cc Sequence 414 AA:

Query Match 59.4%; Score 1409; DB 20; Length 414;
 Post Local Similarity 87.8%; Prod No 5 40 144;
 Matches 273; Conservative 16; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNTSIPSNVNLGCGRGKRTSPNVLWPTGHSPLVVRHQLSHLYAEFGSPWCEARSL 60
 DB 73 MNTSVSVSTNLEGGVGRQVSPNVLWPTGHSPLVVRHQLSHLYAEFGSPWCEARSL 132
 QY 61 EHLLVRRKELKKRVSNGPVSPTGSKRPAHPAVNSIVASGYHGVMSCEGCKAF 120
 DB 133 EHLLVRRKELKKRVSNGPVSPTGSKRPAHPAVNSIVASGYHGVMSCEGCKAF 192
 QY 121 KRSTGHNQYTGATNQTGDKNRKRSQVAPLPRVEYGVWVCSPEEPGYPLVRQR 180
 DB 193 KRSTGHNQYTGATNQTGDKNRKRSQVAPLPRVEYGVWVCSPEEPGYPLVRQR 213
 QY 181 SADEQDLGACAKKRSCHAPRVRELLDALSPQVLTLLFAPPHVLTLSRSAPETAS 240
 DB 214 SASQVGGTNRKAKRTSHTPRVRELLDALSPQVLTLLFAPPHVLTLSRSAPETAS 273
 QY 241 MMSSTKLAKRELVIHMSAKKIPGVVELSTFGVYHLEFSEWMEVIMMGLMKRSIDHCK 300
 DB 274 MMSSTKLAKRELVIHMSAKKIPGVVELSTFGVYHLEFSEWMEVIMMGLMKRSIDHCK 333
 QY 401 LIFAPDLVLDKDEKVEGILLLEPMLATTSPEPKLKQIKRYLVKAMILLNSSTYL 460
 DB 434 LIFAPDLVLDKDEKVEGILLLEPMLATTSPEPKLKQIKRYLVKAMILLNSSTYL 493

RESULT 28

AAV04436 standard; Protein: 740 AA.

AC AAY04446:
 XX 25 JUN-1999 (first entry)
 OT Rat ER beta-4 clone protein sequence.
 DE Estrogen receptor beta: oestrogen receptor beta: ER beta.
 KW Estrogen receptor beta: oestrogen receptor beta: ER beta.
 XX
 OS Rattus sp.
 XX W09905171-A1.
 XX W09905171-A1.
 XX 04-FEB-1999.
 XX 28-JUL-1998: 98W0-0515540.
 XX 30-JUL-1997: 9708-0054210.
 XX 28-JUL-1997: 9708-0053869.
 XX (UYYA) UNIV YALE.
 XX Baron R, Chambon P, Beonten R, Dupont S, Garnier J;
 XX WPI: 1999-14285/12.
 PT Novel murine oestrogen receptor beta genes and splice variants -
 PT useful for treating oestrogen receptor-beta mediated disorders
 XX
 PS Claim 9: Fig 2c: 81pp: English.
 XX The present invention describes the murine complete oestrogen receptor
 CC beta (mER beta c), also designated mER beta-3). The present sequence
 CC represents the rat isoform ER beta-4 of mER beta c. The protein can
 CC be used to screen for agents that modulate or block the interaction
 CC between the ER beta c protein and physiological, stilbene or
 CC triphenylmethylene (anti)estrogens. The antibody which binds the protein
 CC can be used in assays to detect activation of the ER beta protein and
 CC complementary nucleic acid can be used to inhibit expression of the
 CC ER beta c nucleic acid.
 XX
 SO Sequence 730 AA:
 Query Match 59.4%; Score 1499; DB 20; Length 730;
 Best Local Similarity 87.8%; Pred. No. 1,2c-133;
 Matches 273; Conservative 16; Mismatches 22; Indels 0; Gaps 0;
 QY 1 MNYSPSVNLEGGGRQTTSPNVLMPTGHLSPVVRQLSHYAEFGKSNWCEARSL 60
 DB 73 MNYSPSVNLEGGGRQTTSPNVLMPTGHLSPVVRQLSHYAEFGKSNWCEARSL 132
 QY 61 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDAHFCVSDVAGYGVWSVEGCKAF 120
 DB 133 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDAHFCVSDVAGYGVWSVEGCKAF 192
 QY 121 KRSLGHNHYTGATNQTITDKNPKRSQACPI KRCYFVGMVKGSSRPERGYPVLRKQ 180
 DB 193 KRSLGHNHYTGATNQTITDKNPKRSQACPI KRCYFVGMVKGSSRPERGYPVLRKQ 252
 QY 181 SADOIHCACAKRSCHAPVRPELIDALSPEDVITLLEAPRVLLSRPAPFTAS 240
 DB 253 SSSSDVHCISKAKNGCHAPVRPELIDALSPEDVITLLEAPRVLLSRPAPFTAS 412
 QY 241 MMSSTKRLAKPIVHMTSMARKTPPEVEI SLFVWLLRSQWMEVLMMLMWSIHHQK 400
 DB 413 MMSSTKRLAKPIVHMTSMARKTPPEVEI SLFVWLLRSQWMEVLMMLMWSIHHQK 472
 QY 401 LIFAPDLVDLR 311
 DB 473 LIFAPDLVDLR 383

AA881600
 ID AA881600 standard; Protein: 587 AA.
 AC AA881600:
 XX 19-SEP-2002 (first entry)
 DE Calman crocodylus oestrogen receptor protein Skg ID No:1.
 XX Calman crocodylus; oestrogen receptor; endocrine disrupter; receptor.
 KW Calman crocodylus.
 XX Calman crocodylus.
 XX W0200248160-A1.
 XX 20-JUN-2002.
 XX 15-NOV-2001: 2001W0-J099996.
 XX 14-DEC-2000: 2000JP-0479994.
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX Sumida K;
 XX WPI: 2002-557614/59.
 XX N-PSDB; A0072931.
 XX
 PS Claim 1: Page 60-63; 81pp: Japanese.
 XX The present sequence represents an oestrogen receptor isolated from
 CC Calman crocodylus. The oestrogen receptor gene can be used in testing
 CC systems for evaluating the ability of chemicals to control the
 CC oestrogen receptor activity, e.g. assay safety of chemicals and
 CC detecting endocrine disrupters in environment and the genes and the
 CC DNA are useful in two-hybrid assays.
 XX
 SO Sequence 587 AA:
 Query Match 48.0%; Score 1211; DB 23; Length 587;
 Best Local Similarity 55.0%; Pred. No. 2,5e-106;
 Matches 246; Conservative 71; Mismatches 106; Indels 24; Gaps 8;
 QY 22 SPNVLMPTGHLSPVVRQLSHYAEFGKSNWCEARSL EHTLPVNRFTLKRKVSNG 78
 DB 100 SPNVLMPTGHLSPVVRQLSHYAEFGKSNWCEARSL EHTLPVNRFTLKRKVSNG 156
 QY 79 RCASPVTPG-----SKRDAHFCVSDVAGYGVWSVEGCKAFKKSIGHNHYTC 133
 DB 157 ELMSSISKASLSMSEIKELRYCAVCDVAGYGVWSVEGCKAFKKSIGHNHYTC 216
 QY 134 ATNQTITDKNPKRSQACPI KRCYFVGMVKGSSRPERGYPVLRKQSLFVLRKQ 193
 DB 217 ATNQTITDKNPKRSQACPI KRCYFVGMVKGSSRPERGYPVLRKQSLFVLRKQ 276
 QY 194 RSCNAPV-----ALSPEDVITLLEAPRVLLSRPAPFTAS 241
 DB 277 TAEKTPPLWTSPLVILKTKKNSVALSLADQVSAALADAPVLYSKYDGRKPNFNASM 436
 QY 242 MMSSTKRLAKPIVHMTSMARKTPPEVEI SLFVWLLRSQWMEVLMMLMWSIHHQK 401
 DB 437 MTLTINLADKELVIMIMAKRVKPVTLTHOVHLEFAMLEILMIGLVKSMREHPQ 496
 QY 402 LFAPIVLDPEKQVETITETMLATTSEPEPI KLGREYLVKAMILLNSMYPV 461
 DB 497 LEAVNLIDPEKQVETITETMLATTSEPEPI KLGREYLVKAMILLNSMYPV 456
 QY 462 TATQVADSSPKLAHLLNAVTDALVWYIAKSLISSQGSRLANLMLSHVHANSKRM 420

457 SSTSLKEKDYHRAVLKLTPTLTHIMAKSGSLGGHBERAQLLTHSHHMKSEKCH 516
4Z1 EHLNMGKKNVVPYDLLEMLNAHV 447
517 EHLNMGKKNVVPYDLLEMLNAHV 543

RESULT 40

AB084791 standard; Protein: 581 AA.

AB084791:

28 AUG 2002 (first entry)

Chemidophorus uniparus oestrogen receptor SHG ID No 1.

Chemidophorus uniparus; oestrogen receptor.

Chemidophorus uniparus.

W0200252010 A1.

04 JUL 2002.

15 NOV 2001; 2001W0-JP09995.

25 DEC 2000; 2000JP-0492262.

(SUMO) SUMITOMO CHEM CO LTD.

Sumida K.

WP1: 2002 500766/53.

N FSD0; AB081466.

Chemidophorus uniparus originated oestrogen receptor genes, applicable in test systems for evaluating ability of chemicals to control oestrogen receptor activity, e.g. assay safety of chemicals and detecting endocrine disruptors.

Claim 1: Page 59 62; 41pp; Japanese.

The invention relates to an oestrogen receptor gene (AB081466) encoding a 581 residue amino acid sequence (AB084791) or its variant, or a sequence encoded by a polynucleotide sequence having not less than 85 % homology with (AB081466) and having oestrogen receptor activity. The genes are applicable in test systems for evaluating the ability of chemicals to control the oestrogen receptor activity, e.g. assay safety of chemicals and detecting endocrine disruptors in environment.

Sequence 581 AA:

Query Match 48.0% Score 1210; DB 23; Length 581.
Best Local Similarity 54.4% Prod. No. 40-106;
Matches 242; Conservative 74; Mismatches 98; Indels 40; Gaps 0.

22 SPNVNLTGPGHSPVAVHPS MVAHPSQWTPASLEHTLVNPEETKPPVSN 78

100 SPVVFQTAFLQSPFHNNQVYVYLFENSSAMKFAFPAFYVDS --ENHBER 155

79 KCAVGVGP GSKFAHCAVAGSYASGYHCVMSFGCAFKERSTGGINDYVAT 135

156 KASNSKSGLSMSETEKTPQAVGNVAVSGYHGVMSFGCAFKERSTGGINDYVAT 215

146 NGPTLDRKRSQVATNARKYEVQWVQVSGFPAVYLVVPGVSAFVQWAFVAP 195

216 NGPTLDRKRSQVATNARKYEVQWVQVSGFPAVYLVVPGVSAFVQWAFVAP 275

196 GCHAPVGE LLLD ALSTEGVLTLLAPPPVLS RPSAG 235

267 GRNAVAVTARNITLWSPIMKRSKNSDALSTAPGVASLITAPPPVAVSVVPS 329

235 EASMMSTKALRLVHM:SWAKTGEVSLHYRLFSQWMLMGIMKRSI 295
427 FESASVMLTLNLAHRLVMTWAKVRFVDAHLYVHLR:AMLEMLMLTWSTL 486
296 DHPRLTAVGVYDRKQVWVLTDLREMLATISREKLGKRELVKAMLLNS 455
487 EHPKRLFAVNLIDRSQMVVEGFVLEMLATISREKLMNQLQEEFVTKSLTNS 446
456 SMPLVAT QDALSSKLAHLNATVDALVWYAKSGISSQGSMLGLANLMSVRL 414
447 GIVFLSSSTRLEKHHRAVLKLTPTLTHIMAKSGSLGGHBERAQLLTHSHH 506
415 ASKGMELNMGKKNVVPYDLLEMLNAHV 447
507 MSKGMELNMGKKNVVPYDLLEMLNAHV 543

RESULT 41

AA078227 standard; Protein: 582 AA.

AA078227:

13 DEC 2001 (first entry)

Lepomis centrarchidae oestrogen receptor SHG ID No 4.

Lepomis centrarchidae.

JP2001197890-A.

24 JUL 2001.

08 NOV 2000; 2000JP-0440097.

09 NOV 1999; 99JP-0418113.

(SUMO) SUMITOMO CHEM CO LTD.

WP1: 2001 609402/70.

N FSD0; AA168126.

New polynucleotide for controlling oestrogen receptor activity comprises the oestrogen receptor gene.

Claim 1: Page 23-24; 34pp; Japanese.

The invention relates to oestrogen receptors and the encoding genes.

The oestrogen receptor genes can be used for evaluating the ability of controlling oestrogen receptor activity of a chemical substance.

Sequence 582 AA:

Query Match 48.0% Score 1205; DB 22; Length 582.
Best Local Similarity 52.0% Prod. No. 440-106;
Matches 254; Conservative 68; Mismatches 98; Indels 65; Gaps 12.

3 YSLP SNVTNLDGGPGRTSPNVNLTGPGHSPVAVHPSQWTPASLEHTLVNPEETKPPVSN 53

41 YSVTLGGGPGHSPVAVHPSQWTPASLEHTLVNPEETKPPVSN 94

54 WTPASLEHTLVNPEETKPPVSN 101

65 YVPSVYVSGVYVPE --ALSHSYVSGSVAQVDFEMAKRPPVAVSD 146

162 YASVAVHGVMSFGCAFKERSTGGINDYVATNARKYEVQWVQVSGFPAVYLVVPGVSAFVQWAFVAP 161

147 YASVAVHGVMSFGCAFKERSTGGINDYVATNARKYEVQWVQVSGFPAVYLVVPGVSAFVQWAFVAP 206

12 VPSVYVSGVYVPE --ALSHSYVSGSVAQVDFEMAKRPPVAVSD 201

QY 10 TNLGGPGRO-----ITSPVNLWPTGHSPLVVRKQLSHVYAPG 50
 DB 76 TGLPYGPGSAANAGSNGIGGPPPLNSVSPSLMLLHPPOJSPFLQPGQGVVYVLENE 135
 QY 51 KSPWC--EARSLEHTI PVNPEETLKPKVSNIPKASPVTPGQ-----SKPANHCAVMSOYA 103
 DB 136 PSQYTVPEAPAPPAVPP---NSINPRQGGHEKRLASTINDGSMAMESAKETRCVACVNOYA 192
 QY 104 SGHYVWVSCFCKAFKPSLGGHNYVTPATWCTTIDNNRPSQVATLKYVGVGVK 163
 DB 193 SGHYVWVSCFCKAFKPSLGGHNYVTPATWCTTIDNNRPSQVATLKYVGVGVK 163
 QY 164 GQSPRPQGVYVLPQPSADEGIFGAKAPRSQGHAPRV-----ELT 207
 DB 253 GQTPKRPQGPQPMKKKQKQPD-----GPGGFGVSGMGMVANIWSPMLTKPSKNLSA 307
 QY 208 DALSPEDVLTLEFAPPHVLSR--PSAFTFASMMMSLTILAEIVHMTSWAKETFG 265
 DB 308 LSLADQMSALIDAEPP--IIYSEYDTPKPSVSMMLITNLADPEIVHMTSWAKETFG 366
 QY 266 FVPLSLFDVQLLESCMENVLMKIMKPSIDHPCKLFAADVYIDRDEKCVGIIETFD 425
 DB 367 FVPLSLFDVQLLESCMENVLMKIMKPSIDHPCKLFAADVYIDRDEKCVGIIETFD 425
 QY 426 MLATTSRPPELKQHEFYGVKAMII NSMTVLTAT QVANDSEKLAHLINVTAT 384
 DB 427 MLATTSRPPELMNIQPEFYVCTKSTII NSCVTFESSTI KSLFEEDGHPALDEITDIL 486
 QY 385 VVWIAKSGISSQVSMKLANLMLLSHVHASKKQHEHLINMKKNVIVYTLLEMLNA 444
 DB 487 IHMAKAGLILQOCHQKADQLLTLSTIRHMSKQHEHLINMKKNVIVYTLLEMLNA 546
 QY 445 HVL 447
 DB 547 HVL 549
 DB 547 HVL 549
 RESULT 35
 AAG84506
 ID AAG84506 standard; Protein: 595 AA.
 AC AAG84506:
 DT 10-SEP-2001 (first entry)
 XX
 DE Human oestrogen receptor alpha protein mutant K103P.
 XX
 KW Ligand dependent transcriptional factor; oestrogen receptor; PR.
 KW glucocorticoid receptor protein; GR; mineralocorticoid receptor protein;
 KW MR; peroxisome proliferator-activated receptor protein; PPAR;
 KW progesterone receptor protein; PR; retinane X receptor protein; RXR;
 KW thyroid hormone receptor protein; TR; vitamin D receptor protein; VDR;
 KW transactivation; Kralpha; breast cancer; mutant; mutleu.
 XX
 OS Homo sapiens.
 XX
 FN W0200142307-A1.
 FD 14-JUN-2001.
 XX
 XX
 DB 01-DEC-2000; 2000WO-JP0853.
 XX
 XX 07-DEC-1999; 99JP-0348022.
 XX 27-DEC-1999; 99JP-0370667.
 XX 07-JUL-2000; 2000JP-0207011.
 XX 21-10-2000; 2000JP-0230508.
 XX 02-AUG-2000; 2000JP-0234053.
 XX 03-AUG-2000; 2000JP-0234560.
 XX 03-AUG-2000; 2000JP-0234561.
 XX 03-AUG-2000; 2000JP-0234563.
 XX
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX

PI Saito K, Ohe N, Satoh H;
 XX
 UK WPL: 2001-36/866/38.
 XX
 PI Ligand dependent transcriptional factors, nucleic acids encoding them
 PT and cells comprising them and a specified reporter gene, useful for
 PT screening agents for the treatment of breast cancer.
 XX
 XX Claim 15; Page 170-174; 276pp; English.
 XX
 CC The present invention relates to ligand dependent transcriptional factors
 CC including oestrogen receptor (ER) alpha and beta protein, glucocorticoid
 CC receptor protein (GR), mineralocorticoid receptor protein (MR),
 CC peroxisome proliferator-activated receptor protein (PPAR), progesterone
 CC receptor protein (PR), retinane X receptor protein (RXR), thyroid hormone
 CC receptor protein (TR) and vitamin D receptor protein (VDR), the nucleic
 CC acids encoding them and cells comprising them and a specified reporter
 CC gene for the ligand dependent transcriptional factor. These proteins are
 CC useful in the modulation of ligand dependent transcriptional factor
 CC activity. The cells, mutant Kralpha and the polynucleotide encoding it
 CC may be used in assays for facilitating analysing an activity for
 CC transactivation of a reporter gene by a test Kralpha, for screening
 CC mutant ligand dependent transcriptional factors for evaluating an
 CC activity for transactivation of a reporter gene by a test Kralpha and/or
 CC for screening a compound useful for treating a disorder of a mutant
 CC Kralpha, especially breast cancer.
 CC
 SO Sequence 595 AA:
 Query Match 47.7%; Score 1204; DB 22; Length 595;
 Best Local Similarity 59.9%; Pval. No. 1.2e-105;
 Matches 246; Conserved 91; Mismatches 162; Indels 54; Gaps 10;

QY 10 TNLGGPGRO-----ITSPVNLWPTGHSPLVVRKQLSHVYAPG 50
 DB 76 TGLPYGPGSAANAGSNGIGGPPPLNSVSPSLMLLHPPOJSPFLQPGQGVVYVLENE 135
 QY 51 KSPWC--EARSLEHTI PVNPEETLKPKVSNIPKASPVTPGQ-----SKPANHCAVMSOYA 103
 DB 136 PSQYTVPEAPAPPAVPP---NSINPRQGGHEKRLASTINDGSMAMESAKETRCVACVNOYA 192
 QY 104 SGHYVWVSCFCKAFKPSLGGHNYVTPATWCTTIDNNRPSQVATLKYVGVGVK 163
 DB 193 SGHYVWVSCFCKAFKPSLGGHNYVTPATWCTTIDNNRPSQVATLKYVGVGVK 163
 QY 164 GQSPRPQGVYVLPQPSADEGIFGAKAPRSQGHAPRV-----ELT 207
 DB 253 GQTPKRPQGPQPMKKKQKQPD-----GPGGFGVSGMGMVANIWSPMLTKPSKNLSA 307
 QY 208 DALSPEDVLTLEFAPPHVLSR--PSAFTFASMMMSLTILAEIVHMTSWAKETFG 265
 DB 308 LSLADQMSALIDAEPP--IIYSEYDTPKPSVSMMLITNLADPEIVHMTSWAKETFG 366
 QY 266 FVPLSLFDVQLLESCMENVLMKIMKPSIDHPCKLFAADVYIDRDEKCVGIIETFD 425
 DB 367 FVPLSLFDVQLLESCMENVLMKIMKPSIDHPCKLFAADVYIDRDEKCVGIIETFD 425
 QY 426 MLATTSRPPELKQHEFYGVKAMII NSMTVLTAT QVANDSEKLAHLINVTAT 384
 DB 427 MLATTSRPPELMNIQPEFYVCTKSTII NSCVTFESSTI KSLFEEDGHPALDEITDIL 486
 QY 385 VVWIAKSGISSQVSMKLANLMLLSHVHASKKQHEHLINMKKNVIVYTLLEMLNA 444
 DB 487 IHMAKAGLILQOCHQKADQLLTLSTIRHMSKQHEHLINMKKNVIVYTLLEMLNA 546
 QY 445 HVL 447
 DB 547 HVL 549
 DB 547 HVL 549
 RESULT 36
 AAG84513
 ID AAG84513 standard; Protein: 595 AA.

XX Claim 6: Page 131-132; 149pp; English.

XX The present sequence is the protein sequence of the human nuclear
XX oestrogen receptor alpha (1). The invention relates to a non-human
XX mutant organism, especially a transgenic mouse, characterized in
XX that at least one cell comprises: (i) a fusion protein formed from
XX a recombinase (Cre), a hinge region (preferably human (1) D hinge),
XX and a modified ligand binding domain (lbd) of a nuclear oestrogen
XX receptor, especially human (1), its fragment or variant, and (ii)
XX one or more genes or DNA sequences of interest belonging to the
XX genome of the organism, into which one or more recombinase sites
XX of the recombinase protein are inserted. The (2) lbd domain is
XX preferably modified by a G521R, G400V, or M53A/L544A mutation.
XX The fusion protein has no inhibitory, or even zero, recombinase
XX activity in the presence of a natural ligand such as oestradiol,
XX but recombinase activity is induced by a small quantity of a
XX synthetic ligand that has antioestrogenic activity, e.g. tamoxifen
XX or 4-hydroxytamoxifen. The mutant organism or its cells, such
XX as epidermal cells, hepatocytes or adipocytes, are useful in
XX carrying out a spatiotemporally controlled site-specific
XX recombination of a DNA sequence of interest in its natural
XX chromatin environment. It is also used in screening of medicaments
XX for pathological conditions associated with an alteration of the
XX expression and/or function of the DNA sequence of interest, such as
XX skin cancer, inflammation, diabetes, alopecia, obesity, or in
XX promoting hepatic regeneration.

XX Sequence: 595 AA:

Query Match: 47.7%; Score 1204; DB 23; Length 595;
Best Local Similarity: 50.7%; Pred. No. 1,2e-105;
Matches: 245; Conservative: 81; Mismatches: 102; Indels: 54; Gaps: 10;

UY 10 TNRKGRGQ-----TTSPNVI WPTPPI SP-VVHPQI SHI YAPQ 50
DB 76 TLPYPGGSAAFSGNSLGGPPPLNSVSPMLIHPPQISPLFGQGVYYLNE 135
UY 51 KSPWC--EAKSLHTLLVNFETLKRVSNRCASPTGCG--SKRAHGCAGSYIA 103
DB 136 PSYTVTFPACPPAPAYP---NSGNPRQGGPPPLASTNKGSMASAKETPPGAGVSNYA 192
UY 104 SGVHTVWSPCTCKAFKRPSTCHNVTTPATNCTTSPNPPSYACPLPTEVZWY 163
DB 193 SGVHTVWSPCTCKAFKRPSTCHNVTTPATNCTTSPNPPSYACPLPTEVZWY 252
UY 164 GNSPPRCGYPIVPPGPSAPFQIHCACAKRSCGATVPY...PIL 227
DB 253 GGIIPKRRGPRPKLKHRRQRPD---GSPGGEVSGAGLWMAANLWPSPLMKRSKNSLA 307
UY 208 DALSPEQVLTLLAEPPHVLISF--PSAPFTASMMSTLTKATKELVHMSMAKRI 265
DB 308 LSLCLADQVMSALLDAEP--LTYSEYDPTRFPSASMMSTLTKATKELVHMSMAKRI 366
UY 266 FVEISLFDGYVLLPSWMEVLMGLMWPSTIDHPGKLI FADVLVDDEKCEVEGILEPD 225
DB 367 FVDITLHGVHTEFCWAFPTLMGIVWPSMFPKCKLIPADNILLPCKCEVCEWETEPD 429
UY 326 MLATTSRRELKIOKREYICAKAMILLNSMYPVTAL--QDADSRLAHILNVTAL 384
DB 427 MLATTSRREPMNIQGFHVGLKSTILENSGYTETSTLKSELEKNIHVVATKLTDL 486
UY 385 WVAIAKSTISQSSMIA NITMLTSHVPSKNSKEMHILNMCKANVVDLLLEMA 444
DB 487 IHLMAKAGT TLQDQHPALADLITISHIPMSKRMFHI YSMCKRNVVLTLLLEMLTA 546
UY 445 HVL 447
DB 547 HRL 549

RESULT 38
AAG84512

UU AAG84512 standard; Protein: 595 AA.

XX AAG84512:

DE 10-SEP-2001 (first entry)

DE Human oestrogen receptor alpha protein mutant K51H.

KW Ligand dependent transcriptional factor, oestrogen receptor; ER;

KW glucocorticoid receptor protein; GR; mineralocorticoid receptor protein;

KW MR; peroxisome proliferator-activated receptor protein; PPAR;

KW progesterone receptor protein; PR; pregnane X receptor protein; PXR;

KW thyroid hormone receptor protein; TR; vitamin D receptor protein; VDR;

XX transactivation; Eralpha; breast cancer; mutant; metelin.

XX Homo sapiens.

XX W0200142307-A1.

XX 14-JUN-2001.

XX 01-DEC-2000; 2000W0-JP08553.

XX 07-DEC-1999; 99JP-0348022.

XX 27-DEC-1999; 99JP-0370667.

XX 07-JUL-2000; 2000JP-0307011.

XX 21-JUL-2000; 2000JP-0220508.

XX 02-APR-2000; 2000JP-0234053.

XX 03-APR-2000; 2000JP-0235460.

XX 03-APR-2000; 2000JP-0235461.

XX 03-APR-2000; 2000JP-0235463.

XX (SUMO) SUMITOMO CHEM CO LTD.

XX Saito K, Ohe N, Satoh H;

XX WPI: 2001-367866/38.

XX Ligand dependent transcriptional factors, nucleic acids encoding them

XX and cells comprising them and a specified reporter gene, useful for

XX screening agents for the treatment of breast cancer.

XX Disclosure: Page 191-194; 276pp; English.

XX The present invention relates to ligand dependent transcriptional factors

XX including oestrogen receptor (ER) alpha and beta protein, glucocorticoid

XX receptor protein (GR), mineralocorticoid receptor protein (MR),

XX peroxisome proliferator-activated receptor protein (PPAR), progesterone

XX receptor protein (PR), pregnane X receptor protein (PXR), thyroid hormone

XX receptor protein (TR) and vitamin D receptor protein (VDR), the nucleic

XX acids encoding them and cells comprising them and a specified reporter

XX gene for the ligand dependent transcriptional factor. These proteins are

XX useful in the modulation of ligand dependent transcriptional factor

XX activity. The cells, mutant Eralpha and the polynucleotide encoding it

XX may be used in assays for qualitatively analysing an activity for

XX transactivation of a reporter gene by a test Eralpha, for screening

XX mutant ligand dependent transcriptional factors, for evaluating an

XX activity for transactivation of a reporter gene by a test Eralpha and/or

XX for screening a compound useful for treating a disorder of a mutant

XX Eralpha, especially breast cancer.

XX Sequence: 595 AA:

Query Match: 47.7%; Score 1202; DB 22; Length 595;
Best Local Similarity: 50.7%; Pred. No. 1.4e-105;
Matches: 245; Conservative: 82; Mismatches: 102; Indels: 54; Gaps: 10;

UY 10 TNRKGRGQ-----TTSPNVI WPTPPI SP-VVHPQI SHI YAPQ 50
DB 76 TLPYPGGSAAFSGNSLGGPPPLNSVSPMLIHPPQISPLFGQGVYYLNE 135
UY 51 KSPWC--EAKSLHTLLVNFETLKRVSNRCASPTGCG--SKRAHGCAGSYIA 103

```

DB 146 PSTYTVKAEAPVAYRPR NSINRKGQGRERLASTNKGSMAMESAKETRYAVNDYA 192
GY 104 SGAHYGVWSFEFGKAFPKRSIGCHNDYTCVAINQCTIDGNRRKSTGACRLRGYGVGMK 163
DB 193 SGHYGVWSFEFGKAFPKRSIGCHNDYTCVAINQCTIDGNRRKSTGACRLRGYGVGMK 252
GY 164 GGRREKGGVRLVYRHSADQDGLHCACAKRSQSHAPVYV-----EELT 207
DB 253 GGHKRDGRGGMKIKHKGQDQ-----GGEGGVGAGDMDHAAANIWSPIMTKRKKNSIA 407
GY 208 DALSPEDVLTLLAEPPHVLISK--PSAFTTASMMKSLTKADKELVHMSAKRIPG 265
DB 408 LSLADGVNSMLDAEPP--LLYSYEDTPRPSASMMGLTINLADRELVHMIINAKRVG 466
GY 266 FVELSLPDGVNLTESGMVYIMGTIMVPSLIDHSGKLFAPADVLIDEGEKVEGILIEPD 425
DB 467 FVGLTLIDGVHLLFCAMLETIMIGVWRSEHNGKILFAPNILLDGNOKVVERGVEIFD 426
GY 426 MLATTSRPRELKQHKRYLACVAMILLANSMTPLVAT--QVADSSKRLAHLLNAVTDAL 484
DB 427 MLATSSRPFMMNIGGEEFVCTKSTILNSVYTPSLSTIKSLERKDIHIVADKLTITL 486
GY 485 VWVIAKSGTSSQGSQSMILANILMLSHVIRASRCKMEHILNMKCRVAVVYDILLPELNA 444
DB 487 IHLMKAKNLTIDGQHRLADQLLLSHIKMSRKMDEHLYSMKCRVAVVYDILLPELNA 546
GY 445 HVL 447
DB 547 HRL 549

```

RESULT 39

AA084507
ID AA084507 standard: protein: 595 AA.

AC AA084507:

DE 10 SEP 2001 (first entry)

XX Human oestrogen receptor alpha protein mutant S490F.

XX Ligand dependent transcriptional factor: oestrogen receptor; ER;
XX glucocorticoid receptor protein; GR; mineralocorticoid receptor protein;
XX MR; peroxisome proliferator-activated receptor protein; PPAR;
XX progesterone receptor protein; PR; vitamin X receptor protein; RX;
XX thyroid hormone receptor protein; TR; vitamin D receptor protein; VDR;
XX transcription; Eralpha; breast cancer; mutant; nucleic.

OS Homo sapiens

UN W0200142407 A1.

PD 14-JUN 2001.

XX 01 DEC 2000; 2000MO-3P08553.

XX 07 DEC 1999; 99DP 0448022.

XX 27 DEC 1999; 99DP 0470667.

XX 07-JUN 2000; 2000JP-0207011.

XX 21 JUN 2000; 2000JP-0207011.

XX 02 AUG 2000; 2000JP-0214053.

XX 03 AUG 2000; 2000JP-0215460.

XX 03 AUG 2000; 2000JP-0215461.

XX 03 AUG 2000; 2000JP-0215464.

XX (SINO) SINOIPMO CHEM CO LTD.

XX Saito K, Ghe N, Saitoh H.

XX WPI: 2001 467866/38.

XX Ligand dependent transcriptional factors, nucleic acids encoding them
XX and cells comprising them and a specified reporter gene, useful for

PI screening agents for the treatment of breast cancer -
XX claim 16; page 174-177; 276pp; English.

XX The present invention relates to ligand dependent transcriptional factors
XX including oestrogen receptor (ER) alpha and beta protein, glucocorticoid
XX receptor protein (GR), mineralocorticoid receptor protein (MR),
XX peroxisome proliferator-activated receptor protein (PPAR), progesterone
XX receptor protein (PR), progesterone X receptor protein (PRX), thyroid hormone
XX receptor protein (TR) and vitamin D receptor protein (VDR). The nucleic
XX acids encoding them and cells comprising them and a specified reporter
XX gene for the ligand dependent transcriptional factor. These proteins are
XX useful in the modulation of ligand dependent transcriptional factor
XX activity. The cells, mutant Eralpha and the polynucleotide encoding it
XX may be used in assays for qualitatively analysing an activity for
XX transcription of a reporter gene by a test Eralpha, for screening
XX mutant ligand dependent transcriptional factors, for evaluating an
XX activity for transcription of a reporter gene by a test Eralpha and/or
XX for screening a compound useful for treating a disorder of a mutant
XX Eralpha, especially breast cancer.

XX Sequence 595 AA:

Query Match 47.6%; Score 1201; DB 22; Length 595;
Best Local Similarity 50.9%; Pred. No. 2; Acc 10%;
Matches 246; Conservative 80; Mismatches 103; Indels 54; Gaps 10;

```

GY 10 INLEGRGPQ-----TTSNVLMFTYHNSP--LVHKGSLHLYAFVQ 50
DB 76 TLLVYGVGSEAAVAGNSMIGAGPPTLNSVSLTLMILHYPVGLSPETGLHGGVYYEENE 145
GY 51 KSIWG FAVSLTEHTLVNRETLKRVYSNRCASVYTRG-----SKRIAHVAVSOYA 103
DB 146 PSTYTVKAEAPVAYRPR--NSINRKGQGRERLASTNKGSMAMESAKETRYAVNDYA 192
GY 104 SGAHYGVWSFEFGKAFPKRSIGCHNDYTCVAINQCTIDGNRRKSTGACRLRGYGVGMK 163
DB 193 SGHYGVWSFEFGKAFPKRSIGCHNDYTCVAINQCTIDGNRRKSTGACRLRGYGVGMK 252
GY 164 GGRREKGGVRLVYRHSADQDGLHCACAKRSQSHAPVYV-----EELT 207
DB 253 GGHKRDGRGGMKIKHKGQDQ-----GGEGGVGAGDMDHAAANIWSPIMTKRKKNSIA 407
GY 208 DALSPEDVLTLLAEPPHVLISK--PSAFTTASMMKSLTKADKELVHMSAKRIPG 265
DB 408 LSLADGVNSMLDAEPP--LLYSYEDTPRPSASMMGLTINLADRELVHMIINAKRVG 466
GY 266 FVELSLPDGVNLTESGMVYIMGTIMVPSLIDHSGKLFAPADVLIDEGEKVEGILIEPD 425
DB 467 FVGLTLIDGVHLLFCAMLETIMIGVWRSEHNGKILFAPNILLDGNOKVVERGVEIFD 426
GY 426 MLATTSRPRELKQHKRYLACVAMILLANSMTPLVAT--QVADSSKRLAHLLNAVTDAL 484
DB 427 MLATSSRPFMMNIGGEEFVCTKSTILNSVYTPSLSTIKSLERKDIHIVADKLTITL 486
GY 485 VWVIAKSGTSSQGSQSMILANILMLSHVIRASRCKMEHILNMKCRVAVVYDILLPELNA 444
DB 487 IHLMKAKNLTIDGQHRLADQLLLSHIKMSRKMDEHLYSMKCRVAVVYDILLPELNA 546
GY 445 HVL 447
DB 547 HRL 549

```

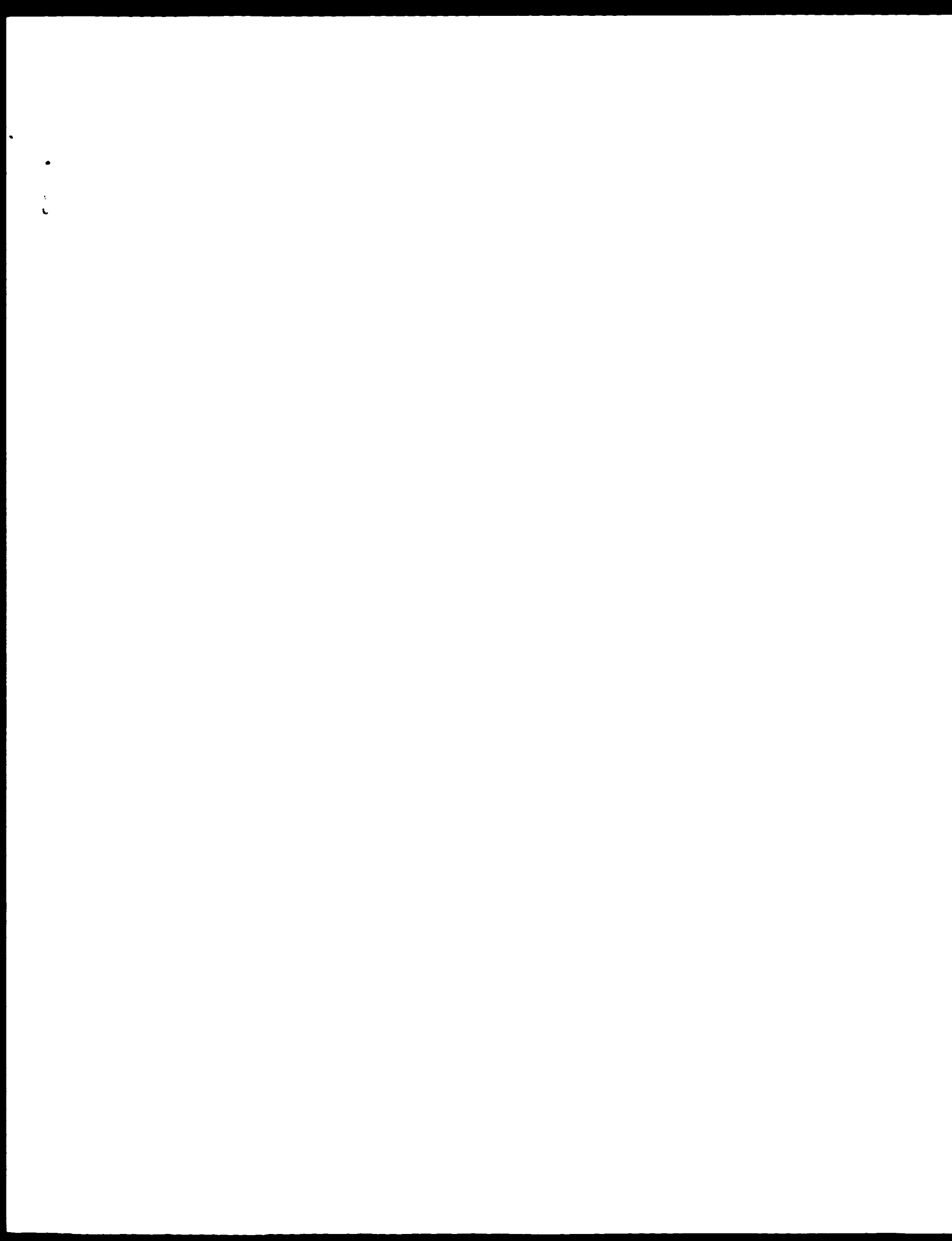
RESULT 40

AA084508
ID AA084508 standard: protein: 595 AA.

AC AA084508:

DE 10 SEP 2001 (first entry)

XX Human oestrogen receptor alpha protein mutant G590D.



101 456 14.1 461 5 PCT US95-15924-2 Sequence 2, Appl 1
 102 456 14.1 461 5 PCT US95-15941-2 Sequence 20, Appl 1
 103 451.5 13.9 764 4 US 09-144-759-20 Sequence 2, Appl 1
 104 450.5 13.9 416 4 US 08-764-870-4 Sequence 4, Appl 1
 105 450.5 13.9 416 4 US 08-980-115-4 Sequence 4, Appl 1
 106 450.5 13.9 797 2 US 08-045-7288-2 Sequence 2, Appl 1
 107 450.5 13.9 797 2 PCT US92-02320A-2 Sequence 2, Appl 1
 108 448.5 13.8 878 4 US 08-653-648A-15 Sequence 15, Appl 1
 109 446.5 13.7 446 2 PCT US95-16411-4 Sequence 4, Appl 1
 110 446.5 13.7 446 5 PCT US95-16411-4 Sequence 4, Appl 1
 111 446 13.7 444 1 US 08-442-411A-4 Sequence 2, Appl 1
 112 444.5 13.7 446 4 US 08-776-844-2 Sequence 18, Appl 1
 113 443.5 13.6 746 4 US 09-144-759-18 Sequence 20, Appl 1
 114 439.5 13.3 440 1 US 08-705-771-20 Sequence 2, Appl 1
 115 435.5 13.3 460 1 US 08-342-411A-2 Sequence 4, Appl 1
 116 425 12.9 323 2 US 09-132-619-4 Sequence 4, Appl 1
 117 425 12.9 323 3 US 09-282-8048-4 Sequence 4, Appl 1
 118 425 12.9 323 4 US 09-510-654-4 Sequence 4, Appl 1
 119 422 12.8 576 4 US 08-653-648A-5 Sequence 5, Appl 1
 120 420 12.7 536 4 US 08-653-648A-13 Sequence 13, Appl 1

ALIGNMENTS

RESULT 1
 US 09-139-617-1
 ? Sequence 1, Application US/09139617
 ? Patent No. 6222015

? GENERAL INFORMATION:
 ? APPLICANT: MILKINSON, HILARY
 ? TITLE OF INVENTION: ESTROGEN RECEPTOR
 ? FILE REFERENCE: 20047Y
 ? CURRENT APPLICATION NUMBER: US/09/139,617
 ? CURRENT FILING DATE: 1998-08-25
 ? EARLIER APPLICATION NUMBER: 60/058,271
 ? EARLIER FILING DATE: 1997-09-08
 ? EARLIER APPLICATION NUMBER: 60/060,520
 ? EARLIER FILING DATE: 1997-09-08
 ? NUMBER OF SEQ ID NOS: 22
 ? SOFTWARE: FastSeq for Windows Version 3.0
 ? SEQ ID NO 1
 ? LENGTH: 548
 ? TYPE: PRT
 ? ORGANISM: HUMAN
 US 09-139-617-1

Query Match 100.0% Score 2522; DB 4; Length 548;
 Best Local Similarity 100.0%; Prod. No. 1.9e-265;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 MNYSTPSNVNINLEGGDQKQTTSPNVLMPTGMLSPILVVRKOLSHLYAEFOKSWCEANSL 60
 DB 72 MNYSTPSNVNINLEGGDQKQTTSPNVLMPTGMLSPILVVRKOLSHLYAEFOKSWCEANSL 131
 UY 61 FHETLVNRETKRKVSNSRPGASVPTGDSKRRAHFAVSDVASGYHYGVMSRPGKAF 120
 DB 132 FHETLVNRETKRKVSNSRPGASVPTGDSKRRAHFAVSDVASGYHYGVMSRPGKAF 191
 UY 121 KRSTGQINDVLTATNCTTDKRRKSCGAPLKYGVYVSWKCSPPREPQYPLVPOR 180
 DB 192 KRSTGQINDVLTATNCTTDKRRKSCGAPLKYGVYVSWKCSPPREPQYPLVPOR 251
 UY 181 SAHQDLCACAKKSGSHAPVRELLDASIPRQVLTLLAEPRVILLSPISAPFTAS 240
 DB 252 SAHQDLCACAKKSGSHAPVRELLDASIPRQVLTLLAEPRVILLSPISAPFTAS 311
 UY 241 MMSSTELAKRELVIHMSWAKIPGVEVLSIFGVVLLSCMMREVIHMSSTIDHRC 300
 DB 312 MMSSTELAKRELVIHMSWAKIPGVEVLSIFGVVLLSCMMREVIHMSSTIDHRC 371
 UY 401 LLEAPVLDVDRKVEVEITLLEPMLATTSPPRELKQREYGVCEKAMILLNSMTL 460
 DB 472 LLEAPVLDVDRKVEVEITLLEPMLATTSPPRELKQREYGVCEKAMILLNSMTL 548

DB 472 LLEAPVLDVDRKVEVEITLLEPMLATTSPPRELKQREYGVCEKAMILLNSMTL 431
 UY 461 VTAQADSSPKLAHLNAVITMAVWVIAKSGISQDSRKAANLMLSHRBASNRGM 420
 DB 432 VTAQADSSPKLAHLNAVITMAVWVIAKSGISQDSRKAANLMLSHRBASNRGM 491
 UY 421 FHLLMMGRKRVVIVYDILLMLNAHVKRKNSTINSDNSVMAHNSKRSRQVSG 477
 DB 492 FHLLMMGRKRVVIVYDILLMLNAHVKRKNSTINSDNSVMAHNSKRSRQVSG 548

RESULT 2
 US-09-561-741A-1
 ? Sequence 1, Application US/09561741A
 ? Patent No. 6458551
 ? GENERAL INFORMATION:
 ? APPLICANT: MILKINSON, HILARY
 ? TITLE OF INVENTION: ESTROGEN RECEPTOR
 ? FILE REFERENCE: 20047Y
 ? CURRENT APPLICATION NUMBER: US/09/561,741A
 ? CURRENT FILING DATE: 2000-04-26
 ? PRIOR APPLICATION NUMBER: 09/139,617
 ? PRIOR FILING DATE: 1998-08-25
 ? PRIOR APPLICATION NUMBER: 60/058,271
 ? PRIOR FILING DATE: 1997-09-08
 ? PRIOR APPLICATION NUMBER: 60/060,520
 ? NUMBER OF SEQ ID NOS: 22
 ? SOFTWARE: FastSeq for Windows Version 3.0
 ? SEQ ID NO 1
 ? LENGTH: 548
 ? TYPE: PRT
 ? ORGANISM: HUMAN
 US 09-561-741A-1

Query Match 100.0% Score 2522; DB 4; Length 548;
 Best Local Similarity 100.0%; Prod. No. 1.9e-265;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 MNYSTPSNVNINLEGGDQKQTTSPNVLMPTGMLSPILVVRKOLSHLYAEFOKSWCEANSL 60
 DB 72 MNYSTPSNVNINLEGGDQKQTTSPNVLMPTGMLSPILVVRKOLSHLYAEFOKSWCEANSL 131
 UY 61 FHETLVNRETKRKVSNSRPGASVPTGDSKRRAHFAVSDVASGYHYGVMSRPGKAF 120
 DB 132 FHETLVNRETKRKVSNSRPGASVPTGDSKRRAHFAVSDVASGYHYGVMSRPGKAF 191
 UY 121 KRSTGQINDVLTATNCTTDKRRKSCGAPLKYGVYVSWKCSPPREPQYPLVPOR 180
 DB 192 KRSTGQINDVLTATNCTTDKRRKSCGAPLKYGVYVSWKCSPPREPQYPLVPOR 251
 UY 181 SAHQDLCACAKKSGSHAPVRELLDASIPRQVLTLLAEPRVILLSPISAPFTAS 240
 DB 252 SAHQDLCACAKKSGSHAPVRELLDASIPRQVLTLLAEPRVILLSPISAPFTAS 311
 UY 241 MMSSTELAKRELVIHMSWAKIPGVEVLSIFGVVLLSCMMREVIHMSSTIDHRC 300
 DB 312 MMSSTELAKRELVIHMSWAKIPGVEVLSIFGVVLLSCMMREVIHMSSTIDHRC 371
 UY 401 LLEAPVLDVDRKVEVEITLLEPMLATTSPPRELKQREYGVCEKAMILLNSMTL 460
 DB 432 VTAQADSSPKLAHLNAVITMAVWVIAKSGISQDSRKAANLMLSHRBASNRGM 491
 UY 421 FHLLMMGRKRVVIVYDILLMLNAHVKRKNSTINSDNSVMAHNSKRSRQVSG 477
 DB 492 FHLLMMGRKRVVIVYDILLMLNAHVKRKNSTINSDNSVMAHNSKRSRQVSG 548

US-08-836-620A-3
Sequence 3, Application US/08846620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/836,620A
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03943
FILING DATE:
APPLICATION NUMBER: GB 9518272.1
FILING DATE: 08-SEP-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9605550.4
FILING DATE: 15-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9607542.0
FILING DATE: 11-APR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9609576.5
FILING DATE: 08-MAY-1996
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-836-620A-3

Query Match: 99.6%; Score 2512; DB 2; Length 485;
Host Local Similarity: 99.8%; Prot. No. 1 90-264;
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYSIPSNVTNLDEGPGKROTTSPNVLMPTGHLSPVVRHQSILYAEPOKSWFKAESL 60
DB 9 MNYSIPSNVTNLDEGPGKROTTSPNVLMPTGHLSPVVRHQSILYAEPOKSWFKAESL 68
QY 61 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDHFAVCSGYHYGWSGEGKAEF 120
DB 69 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDHFAVCSGYHYGWSGEGKAEF 128
QY 121 KRSTGGHNDYTCPATNCTTDKNRPKSGQACPLKCYVGVGVKCGSPREPGYPLVPPGR 180
DB 129 KRSTGGHNDYTCPATNCTTDKNRPKSGQACPLKCYVGVGVKCGSPREPGYPLVPPGR 188
QY 181 SADBQJHCAKAKSSSHAPVPELLDALSPQLVLTLEAEPPHVLISPSAPFTEAS 240
DB 189 SADBQJHCAKAKSSSHAPVPELLDALSPQLVLTLEAEPPHVLISPSAPFTEAS 248
QY 241 MMSSTIKLAKELVHMLSMWKKIPGEVELSLFVQVRLIESQWMEVLMMSIHHPPK 300
DB 249 MMSSTIKLAKELVHMLSMWKKIPGEVELSLFVQVRLIESQWMEVLMMSIHHPPK 308
QY 401 LIFAPDVLDPPEKCYVPELLETFDMLATTSPEPELKQHKREYLVKAMILLNSMYPL 360
DB 409 LIFAPDVLDPPEKCYVPELLETFDMLATTSPEPELKQHKREYLVKAMILLNSMYPL 368
QY 461 VTATQDADSSKRLAHLLNAVTDALVWYIAKSGISSQSGSMPLANLMLLSHVHANSKCM 420
DB 469 VTATQDADSSKRLAHLLNAVTDALVWYIAKSGISSQSGSMPLANLMLLSHVHANSKCM 428
QY 421 EHLNMMCKNVVPPYDILLMLNNAVJLGGKSSITSGSPADSKSKESQVOSQ 477
DB 429 EHLNMMCKNVVPPYDILLMLNNAVJLGGKSSITSGSPADSKSKESQVOSQ 485

RESULT 4
US-08-836-620A-2
Sequence 2, Application US/08846620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/836,620A
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/03943
FILING DATE:
APPLICATION NUMBER: GB 9518272.1
FILING DATE: 08-SEP-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9605550.4
FILING DATE: 15-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9607542.0
FILING DATE: 11-APR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9609576.5
FILING DATE: 08-MAY-1996
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 485 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US-08-836-620A-2

Query Match: 89.2%; Score 2249; DB 2; Length 485;
Host Local Similarity: 88.5%; Prot. No. 7 76-236;
Matches 422; Conservative 23; Mismatches 32; Indels 0; Gaps 0;

QY 1 MNYSIPSNVTNLDEGPGKROTTSPNVLMPTGHLSPVVRHQSILYAEPOKSWFKAESL 60
DB 9 MNYSIPSNVTNLDEGPGKROTTSPNVLMPTGHLSPVVRHQSILYAEPOKSWFKAESL 68
QY 61 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDHFAVCSGYHYGWSGEGKAEF 120
DB 69 EHTLPVNRFTLKRKVSNGRCASPVTPGSKRDHFAVCSGYHYGWSGEGKAEF 128
QY 121 KRSTGGHNDYTCPATNCTTDKNRPKSGQACPLKCYVGVGVKCGSPREPGYPLVPPGR 180
DB 129 KRSTGGHNDYTCPATNCTTDKNRPKSGQACPLKCYVGVGVKCGSPREPGYPLVPPGR 188
QY 181 SADBQJHCAKAKSSSHAPVPELLDALSPQLVLTLEAEPPHVLISPSAPFTEAS 240
DB 189 SADBQJHCAKAKSSSHAPVPELLDALSPQLVLTLEAEPPHVLISPSAPFTEAS 248
QY 241 MMSSTIKLAKELVHMLSMWKKIPGEVELSLFVQVRLIESQWMEVLMMSIHHPPK 300
DB 249 MMSSTIKLAKELVHMLSMWKKIPGEVELSLFVQVRLIESQWMEVLMMSIHHPPK 308
QY 401 LIFAPDVLDPPEKCYVPELLETFDMLATTSPEPELKQHKREYLVKAMILLNSMYPL 360
DB 409 LIFAPDVLDPPEKCYVPELLETFDMLATTSPEPELKQHKREYLVKAMILLNSMYPL 368
QY 461 VTATQDADSSKRLAHLLNAVTDALVWYIAKSGISSQSGSMPLANLMLLSHVHANSKCM 420
DB 469 VTATQDADSSKRLAHLLNAVTDALVWYIAKSGISSQSGSMPLANLMLLSHVHANSKCM 428

Db 409 LIFAPDLVLRDEKCEVESILETFIMLAITAFPELKIQHKREYLVKAMILLNSMYHL 468
QY 461 VTATQADSSKRLAHILNAVTDALVWIAKSGISSQOQSMRIANLMLSHVRHNSNM 420
Db 469 ATASQAFSSSKRLTHIINAVTDALVWIAKSPISSQOQSVRIANLMLSHVRHNSNM 428
QY 421 EHILNMCKRVVYVYIILEMNAHVIKCKSSITSGGCCSTEDSKSKSGSONLQSO 477
Db 429 EHILSMCKRVVYVYIILEMNAHVIKCKSSITSGGCCSTEDSKSKSGSONLQSO 485

RESULT 7
US-08-836-620A-14
Sequence 14, Application US/08836620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPI)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08-836-620A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/03933
FILING DATE:
APPLICATION NUMBER: GB 9518272.1
FILING DATE: 08-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9605550.4
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9607532.0
FILING DATE: 11-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9609576.5
FILING DATE: 08-MAY-1996
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 484 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-836-620A-14

Query Match: 88.2%; Score 2255; DB 2; Length 484,
Best Local Similarity: 88.0%; Prod. No. 3, 1e-213;
Matches 419; Conservative 23; Mismatches 34; Indels 0; Gaps 0;

QY 1 MNTSPNSVNTNFGPSPGPTTSPNVLMPTGSHLSPIVVRHQLSHLYAEQKSIWYAPSL 60
Db 9 MNYSPSSSTNLEGGPQROTASPNVLMPTSGHLSPLAIGHQSLYAEQKSIWYAPSL 68
QY 61 FHTI PVNPFTH KPKVSNPCASPVYTGSGKPDHAPCAVCSDAVSGHYGVWVSCGYKAF 120
Db 69 FHTI PVNPFTH KKKLGSGCASPVYTSIRKDHAPCAVCSDAVSGHYGVWVSCGYKAF 128
QY 121 KRSIQGHNVYIYATNACITIGNPPSSVA*74 PACTVWVWVWVWVWVWVWVWVWVWV 180
Db 129 KRSIQGHNVYIYATNACITIDKNPKKQCGRLKCYEVMWVWVWVWVWVWVWVWVWVWV 188
QY 181 SADIOLHGCAGKAKKSGSHARVRELLLDALSPQVLVTLGAPRPVILSRSPAPTEAS 240
Db 189 SASDVHGLNKAARTSGHTPRVKLLINLSLPEQLVITLLEAPPVVLVSRSPAPTEAS 248
QY 241 MMSLSTKLAKELVHMIISMAKKINGVEVSLFVQVRLDESGWMEVIMMGIWMSIDHCK 400
Db 249 MMSLSTKLAKELVHMIISMAKKINGVEVSLFVQVRLDESGWMEVIMMGIWMSIDHCK 408

QY 401 LIFAPDLVLRDEKCEVESILETFIMLAITSPPELKIQHKREYLVKAMILLNSMYHL 460
Db 409 LIFAPDLVLRDEKCEVESILETFIMLAITAFPELKIQHKREYLVKAMILLNSMYHL 468
QY 461 VTATQADSSKRLAHILNAVTDALVWIAKSGISSQOQSMRIANLMLSHVRHNSNM 420
Db 469 ATASQAFSSSKRLTHIINAVTDALVWIAKSPISSQOQSVRIANLMLSHVRHNSNM 428
QY 421 EHILNMCKRVVYVYIILEMNAHVIKCKSSITSGGCCSTEDSKSKSGSONLQSO 476
Db 429 EHILSMCKRVVYVYIILEMNAHVIKCKSSITSGGCCSTEDSKSKSGSONLQSO 484

RESULT 8
US-08-836-620A-15
Sequence 15, Application US/08836620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPI)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08-836-620A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/03933
FILING DATE:
APPLICATION NUMBER: GB 9518272.1
FILING DATE: 08-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9605550.4
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9607532.0
FILING DATE: 11-APR-1996
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-836-620A-15

Query Match: 80.0%; Score 2017; DB 2; Length 484,
Best Local Similarity: 100.0%; Prod. No. 9, 2e-211;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LSPVVRHQLSHLYAEQKSIWYAPSLFHTI PVNPFTH KPKVSNPCASPVYTGSGKPD 92
Db 2 LSPVVRHQLSHLYAEQKSIWYAPSLFHTI PVNPFTH KPKVSNPCASPVYTGSGKPD 61
QY 93 AHFVAVSDVYASGVWVSCGYKAFKPSIQSHINYICPATNCTIIRKPKKSNQACR 152
Db 62 AHFVAVSDVYASGVWVSCGYKAFKPSIQSHINYICPATNCTIIRKPKKSNQACR 121
QY 153 LKRYEVMWV 212
Db 122 LKRYEVMWV 181
QY 213 EQLVITLLEAPRPVILSRSPAPTEASMMSLKLAKELVHMIISMAKKINGVEVSLF 272
Db 182 EQLVITLLEAPRPVILSRSPAPTEASMMSLKLAKELVHMIISMAKKINGVEVSLF 241


```

7 21P:20005
8
9 COMPUTER READABLE FORM:
10
11 METHOD TYPE: floppy disk
12
13 COMPUTER: IBM PC compatible
14
15 COMMENTS: SYSTEM: 1041 DOS 2.05, DOS

```



```

RESULT 14
US-08-846-620A-16
? Sequence 16, Application US/08/4620A
? Patent No. 5958710
? GENERAL INFORMATION:
? APPLICANT:
? TITLE OF INVENTION: orphan receptor
? NUMBER OF SEQUENCES: 19
? COMPUTER READABLE FORM:
? MEDIUM TYPE: floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25 (PRO)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/46.620A
? FILING DATE:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/JP96/03933
? FILING DATE:
? APPLICATION NUMBER: GR 9518272.1
? FILING DATE: 08-SEP-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GR 9605550.4
? FILING DATE: 15-MAR-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GR 9607532.0
? FILING DATE: 11-APR-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: GB 9609576.5
? FILING DATE: 08-MAY-1996
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 596 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? ORIGINAL SOURCE:
? ORGANISM: Rattus rattus
? US-08-846-620A 16

Query Match 47.94; Score 1192.5; DB 2; Length 596;
Host Local Similarity 48.84; Pred. No. 9,70-121;
Matches 243; Conservative 91; Mismatches 104; Indels 55; Gaps 11;

QY 15 GPGGQ-----TTSPNVLWPTPGHSPV---VHQGSHIYARQGP 63
DB 86 GPGSAAPAFANSGAFAPQNSVSPSPMTLHPHPHSPPIHPCGIVYVYIFNEP-----141
PB

QY 54 WGEARSLFHTP---VNPRTIKRVSQNPASFTVDFQ-----SKRIAHFQAVSGYASG 105
DB 142 SAYAAVDITGPAPFVPSNSQNPQNCPPLSSSPKQNMIMESAKETPYCAVCNDYASG 199
PB

QY 106 YHYGVMSFGCKKAFKRSIQGHNDYIGPATNCTIKKPKRSQVAPLPGVYVWVKQ 165
DB 200 YHYGVMSFGCKKAFKRSIQGHNDYIGPATNCTIKKPKRSQVAPLPGVYVWVKQ 259
PB

QY 166 SPPEPVYVIVPQPSAHE-----QIHQAGKAPKSG-----CHAPVYRELLDANSP 212
DB 260 TPKSPPTPMKRRFGGPILECFHEHETCTMFAVLLNCTLVYHTEKRSFAL--STVA 217
PB

QY 213 EQVLTLTFAPRPVLLS--KPSAFTFASMMSLIKAIKELVHPISMKKIGVEHSL 271
DB 418 DQMSVALIDAPRPILVSPYDPSPESSMMGLINLADRELVHMIINMARVGFQDLNL 477
PB

QY 272 FDOVLLFSWMEVIMM--MMKSLIHFGKLFIAEDVILPEEGVEGGLIEFEMLAT 331
DB 478 HDQVHLIFCAWLETIMIGVWRSMEHSGKLLFANILLDQNGKCVGSGWETDMLATS 437
PB

QY 332 SRPRLKIQRFYIVCAAMILLNSMPTLVAT--QDADSSPKLAHLINAVTALVWVIAK 490
DB 438 SRPRMNIQDEEFGKSLIILNSGYTFLSLKSLFEEDHILHVLDKINDTLIDIMAK 497
PB

QY 491 SGISGQSGMPLANIMLSVPHASNKGMFHI--NPKCVWVIVVYV--IIMNAVYKSG 450
PB

```

```

DB 498 AGLLQGHRELALILLSHIHRMSKGMELIYNKCKNVVITVILLFMAIRIAP 557
QY 451 KSLTIQSFQSPAERSSK 468
DB 558 ASRM---GAPPEERSQD 572

RESULT 15
US-08-826-361a-5
? Patent No. 5223606
? APPLICANT: BLAUDIN DE THE, HUGHES, MARCILLO, AGNES/TIGLATS,
? PIERRE/DEJAN, ANNE
? TITLE OF INVENTION: STEROID/HYROID HORMONE RECEPTOR-RELATED
? PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HERATOCELLULAR CARCINOMA
? NUMBER OF SEQUENCES: 11
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/144,140
? FILING DATE: 17-DEC-1987
? PRIOR APPLICATION DATA:
? SEQ ID NO:5;
? LENGTH: 410
? 5223606-5

Query Match 46.08; Score 1160; DB 6; Length 410;
Host Local Similarity 59.08; Pred. No. 1,80-117;
Matches 219; Conservative 70; Mismatches 56; Indels 26; Gaps 5;

QY 96 CAWSDYASGVHYGVMSFGCKKAFKRSIQGHNDYIGPATNCTIKKPKRSQVAPLPGK 155
DB 1 CAWINDYASGVHYGVMSFGCKKAFKRSIQGHNDYIGPATNCTIKKPKRSQVAPLPGK 50
QY 156 GYFVWVWVGSPRPVYVIVPQPSAHE-----GEQKQGVNSAGMPPANIMPSPIMK 203
DB 61 GYFVWVWVGSPRPVYVIVPQPSAHE-----GEQKQGVNSAGMPPANIMPSPIMK 115
QY 204 ----ELLDAISPEQVITLLEAPRPVLLSR--PSAPTEASMMSLIKAIKELVHMI 257
DB 116 KSKKNSLALSLTADQMSALLDAP--IYSEDPTRPEASMMGLINLADRELVHMI 173
QY 258 SWAKKIGVEHSLFDOVLLFSWMEVIMMGLMMSIDPQGLFAPRPVILDRQGVK 317
DB 174 MWARVPSFVITLHPVHLLCAWLETIMIGVWRSMEHPIVAPNILLDQNGKCV 233
QY 317 INAVTALVWVIAKSGISQDQSMKLANILLSHVHASNKGMELIYNKCKNVVIVYD 436
DB 294 LEKLTIDILHIMAKAGLILCOOHORLAQILLSHIHRMSKGMELIYNKCKNVVYD 353
QY 437 LILKELNHYL 447
DB 454 LILKELNHYL 364

RESULT 16
US-09-249-645-1
? Sequence 1, Application US/09/249645
? Patent No. 6228990
? GENERAL INFORMATION:
? APPLICANT: Lundgren, Jan
? APPLICANT: Thorsell, Ann-Gerd
? APPLICANT: Engstrom, Owe
? APPLICANT: Bonn, Tomas
? APPLICANT: Carlquist, Mats
? APPLICANT: Brzozowski, Andrzej
? APPLICANT: Pike, Ashley
? APPLICANT: Hubbard, Rogerick
? TITLE OF INVENTION: ESTROGEN RECEPTOR LIGANDS
? FILE REFERENCE: 101987-100
? SEQUENCE AFFECTING: 8 NMPBS US/09-249,645
? CURRENT FILING DATE: 1999-02-12

```

PRIOR APPLICATION NUMBER: GB 980 4062.0
 PRIOR FILING DATE: 1998-02-13
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: FASTSPC for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 229
 TYPE: PRT
 ORGANISM: Human
 US 09 249 645 1

Query Match 44.7% Score 1126.5; DB 4; Length 229;
 Best Local Similarity 95.0% Pctid No. 41 to 114;
 Matches 228; Conservative 0; Mismatches 1; Indels 11; Gaps 2;

408 DAISSPGLVITLFAPIHVLISPSAPTEASMMSTIKLAKELVIMISMAKIPGV 267
 1 DAISSPGLVITLFAPIHVLISPSAPTEASMMSTIKLAKELVIMISMAKIPGV 55
 408 PLSIVVPLIESWMEVIMMSTIDHCKLFAPIVILADHCKVVEVILIDHCK 427
 56 PLSIVVPLIESWMEVIMMSTIDHCKLFAPIVILADHCKVVEVILIDHCK 115
 428 LATSREPEIKLGRFELCVKAMILNNSMYPVITVAGUADSSREKLAINVTDAVAV 487
 116 LATSREPEIKLGRFELCVKAMILNNSMYPVITVAGUADSSREKLAINVTDAVAV 169
 408 IAKSSTSSGGSMLANILMLSHVPSNKNRMEHILNMRGVNVI-VYVILLERMAHVL 447
 170 IAKSSTSSGGSMLANILMLSHVPSNKNRMEHILNMRGVNVI-VYVILLERMAHVL 229

RESULT 17

US 09 249 645 2

Sequence 2; Application US/09249645
 Patent No. 6,286,990
 GENERAL INFORMATION:
 APPLICANT: Immunogen, Inc.
 APPLICANT: Thorsell, Ann-Gerd
 APPLICANT: Boquistom, Owe
 APPLICANT: Bond, Tomas
 APPLICANT: Carlqvist, Mats
 APPLICANT: Burzowski, Andrzej
 APPLICANT: Piko, Ashley
 APPLICANT: Hubbard, Frederick
 TITLE OF INVENTION: ESTROGEN RECEPTOR LIGANDS
 FILE REFERENCE: 101967 100
 CURRENT APPLICATION NUMBER: US/09/249, 645
 CURRENT FILING DATE: 1999-02-12
 PRIOR APPLICATION NUMBER: GB 980 4062.0
 PRIOR FILING DATE: 1998-02-13
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: FastSPC for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 228
 TYPE: PRT
 ORGANISM: Rat
 US 09 249 645 2

Query Match 42.2% Score 1064; DB 4; Length 228;

Best Local Similarity 95.2% Pctid No. 26 107;
 Matches 212; Conservative 8; Mismatches 7; Indels 8; Gaps 1;

416 LSPGVVITLFAPIHVLISPSAPTEASMMSTIKLAKELVIMISMAKIPGV 269
 2 LSPGVVITLFAPIHVLISPSAPTEASMMSTIKLAKELVIMISMAKIPGV 61
 470 SLFVGVILSSWMEVIMMSTIDHCKLFAPIVILADHCKVVEVILIDHCK 429
 11 SLFVGVILSSWMEVIMMSTIDHCKLFAPIVILADHCKVVEVILIDHCK 121
 62 SLFVGVILSSWMEVIMMSTIDHCKLFAPIVILADHCKVVEVILIDHCK 121
 440 TTSPEPEIKLGRFELCVKAMILNNSMYPVITVAGUADSSREKLAINVTDAVAV 489
 122 TTSPEPEIKLGRFELCVKAMILNNSMYPVITVAGUADSSREKLAINVTDAVAV 181

490 KSGISSGGSMLANILMLSHVPSNKNRMEHILNMRGVNVI-VYVILLERMAHVL 444
 102 KSGISSGGSMLANILMLSHVPSNKNRMEHILNMRGVNVI-VYVILLERMAHVL 228

RESULT 18

US 08-846-620A 7

Sequence 7; Application US/08846620A

Patent No. 6,958,710

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: orphan receptor

NUMBER OF SEQUENCES: 19

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSPC for Windows #1.0, Version #1.25 (PCT)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/846, 620A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP96/03933

FILING DATE:

APPLICATION NUMBER: GB 9518272.1

FILING DATE: 08 SEP-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9605550.4

FILING DATE: 15 MAR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9607542.0

FILING DATE: 11 APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9609576.5

FILING DATE: 08 MAY-1996

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 226 amino acids

TYPE: amino acid

TOPOLOGY: linear

ORIGINAL SOURCE:

ORGANISM: Rattus rattus

US 08-846-620A 7

Query Match 41.8% Score 1055; DB 2; Length 226;

Best Local Similarity 91.2% Pctid No. 186-106;

Matches 206; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

452 ELVHMSWARKIPVEVELSLIDVYKLLSSWMEVIMMSTIDHCKLFAPIVILADHCK 411
 1 ELVHMSWARKIPVEVELSLIDVYKLLSSWMEVIMMSTIDHCKLFAPIVILADHCK 60
 412 DEGRVGLILETFPIMLATTSREPEIKLGRFELCVKAMILNNSMYPVITVAGUADSSR 471
 61 DEGRVGLILETFPIMLATTSREPEIKLGRFELCVKAMILNNSMYPVITVAGUADSSR 120
 472 KLAIHNAVTDVAVWIAKSGTSSGGSMLANILMLSHVPSNKNRMEHILNMRGVN 441
 121 KLAIHNAVTDVAVWIAKSGTSSGGSMLANILMLSHVPSNKNRMEHILNMRGVN 180
 492 VYVILLERMAHVLNMRGVNVI-VYVILLERMAHVLNMRGVNVI-VYVILLERMAHVL 477
 181 VYVILLERMAHVLNMRGVNVI-VYVILLERMAHVLNMRGVNVI-VYVILLERMAHVL 226

RESULT 19

US 08-564-264 1

Sequence 1; Application US/08564264

Patent No. 6,004,600

GENERAL INFORMATION:

APPLICANT: STEWART, Francis

TITLE OF INVENTION: REGULATION OF SITE SPECIFIC


```

1  APPLICATION NUMBER:  CH 960518272
2  FILING DATE: 08 SEP 1995
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER:  CH 9605550.4
5  FILING DATE: 15 MAR 1996
6  PRIOR APPLICATION DATA:
7  APPLICATION NUMBER:  CH 9607542.C
8  FILING DATE: 11 APR 1996
9  PRIOR APPLICATION DATA:

```


QY 182 ADEQLHACKAKRSQGHAVRVELLDALSPQOL-VITLLAEPPHVL----- 229
DB 277 KIDGEGCAGCA-----PREMPDRILLEAF-----LAEVQKSDGQVGGP 314
QY 230 -----SPSAPPTASMMMSLTKLAWEPIVWMLSWAKIPGVEHSLFDVRLTESQW 282
DB 415 GGTGGSSSPNDPVT-----NIOQAARKQILFTIVMAKRIPESSSLPLDQVILLRQW 368
QY 283 MEVIMMOLMRSIDHPCIKFAPDVLIDPEKCVGILIEPDLIA-TTSPPELKIQH 341
DB 369 NELLIASFSRSDIVPGCILLALGIVHFNASHA-CVGAIFDPVITELVSKMIPMIMK 427
QY 342 KEVLCKAMILLNSSMRYPLVATQVADSSEKLAHLINAVITAIWVIAVIAKSGISSQOSMR 401
DB 428 TELGCFRALLIFPDQKGL-----SNPSEVEVLERKVASLETYCKQKYPBOOG--K 477
QY 402 LANIMLISHVRHASKMKMEHLIMCKKVVVPPYDILLEMNA 444
DB 478 FAKILLRLPALRSIGIKLEHLFPFKLIGDTPIDTIFLMEMLFA 520

RESULT 34

US-08-980-115-7
Sequence 7, Application US/08980115
Patent No. 625622
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S.
APPLICANT: Hackett, John D.
APPLICANT: Fletchrick, Robert J.
APPLICANT: Wagner, Richard L.
APPLICANT: Kushner, Peter J.
APPLICANT: Apiletti, James W.
APPLICANT: West, Brian L.
APPLICANT: Shiau, Andrew K.
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
FILE REFERENCE: OCAI-246/0205
CURRENT APPLICATION NUMBER: US/08/980,115
EARLIER FILING DATE: 1997-11-26
EARLIER APPLICATION NUMBER: 08/764,870
EARLIER FILING DATE: 1996-12-13
EARLIER APPLICATION NUMBER: 60/008,606
EARLIER FILING DATE: 1995-12-14
EARLIER APPLICATION NUMBER: 60/008,543
EARLIER FILING DATE: 1995-12-13
EARLIER APPLICATION NUMBER: 60/008,544
EARLIER FILING DATE: 1995-12-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 7
LENGTH: 525
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: DOMAIN
LOCATION: (292)..(523)
OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-7

Query Match 18.3% Score 460.5; DB 4; Length 525;
Best Local Similarity 30.2% Pred. No. 3,5e-41;
Matches 140; Conservative 51; Mismatches 165; Indels 97; Gaps 16.

QY 9 VTNIIEGPPEQITSPNVIWPTPCHISPLVHPEQLSHLVAPKSPWCEARSLHHITPVN- 67
DB 128 ISSMSGSPGPPAP-----PG-----FSGPVSSPQINSTV---SLPQGG 164
QY 68 KEIKKRVSSNR CASVITGSSKRIAHAVAVSIVASGVHYVWVSSPPCKATFK 121
DB 165 SGPEIVKPPVIAVPIGHPPEPPGQACK--PLCAICGHRSSGHEVGVYSSEGVKQKPK 222
QY 122 PSIECHNIYCPATNPTTIKKKPKSGQANPKKPVAVWVQSSPEPFGVPIVQKES 161

DB 223 PIIPEHLLSSPECHPEVITPEYSAHEQZYTHYKCLAIEMPEFVAPFPE----- 276
QY 182 ADEQLHACKAKRSQGHAVRVELLDALSPQOL-VITLLAEPPHVL----- 229
DB 277 KIDGEGCAGCA-----PREMPDRILLEAF-----LAEVQKSDGQVGGP 314
QY 230 -----SPSAPPTASMMMSLTKLAWEPIVWMLSWAKIPGVEHSLFDVRLTESQW 282
DB 415 GGTGGSSSPNDPVT-----NIOQAARKQILFTIVMAKRIPESSSLPLDQVILLRQW 368
QY 283 MEVIMMOLMRSIDHPCIKFAPDVLIDPEKCVGILIEPDLIA-TTSPPELKIQH 341
DB 369 NELLIASFSRSDIVPGCILLALGIVHFNASHA-CVGAIFDPVITELVSKMIPMIMK 427
QY 342 KEVLCKAMILLNSSMRYPLVATQVADSSEKLAHLINAVITAIWVIAVIAKSGISSQOSMR 401
DB 428 TELGCFRALLIFPDQKGL-----SNPSEVEVLERKVASLETYCKQKYPBOOG--K 477
QY 402 LANIMLISHVRHASKMKMEHLIMCKKVVVPPYDILLEMNA 444
DB 478 FAKILLRLPALRSIGIKLEHLFPFKLIGDTPIDTIFLMEMLFA 520

RESULT 34

US-07-952-800-2
Sequence 2, Application US/07952800
Patent No. 540925
GENERAL INFORMATION:
APPLICANT: OZARO, KEIKO
TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR RECEPTOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND AND TOWNSEND
STREET: 1 MARKET PLAZA, STEWART TOWER, SUITE 2000
CITY: SAN FRANCISCO
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
RELION TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT NUMBER: 05/007,952,800
FILING DATE: 19920928
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A
REGISTRATION NUMBER: 31,677
REFERENCE TO PARENT NUMBER: 15280-21-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NOS: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US 07 952-800-2

Query Match 18.2% Score 458.5; DB 1; Length 533;
Best Local Similarity 30.2% Pred. No. 5,9e-41;
Matches 140; Conservative 63; Mismatches 163; Indels 97; Gaps 17;

QY 9 VTNIIEGPPEQITSPNVIWPTPCHISPLVHPEQLSHLVAPKSPWCEARSLHHITPVN- 67
DB 136 ISSMSGSPGPPAP-----PG-----FSGPVSSPQINSTV---SLPQGG 172
QY 68 KEIKKRVSSNR CASVITGSSKRIAHAVAVSIVASGVHYVWVSSPPCKATFK 121

Db 798 CTIMMOLPOEVLKAGVSGFELPMKVLILNLIPLLELSLQFHEEMSSYIPEL1KA 456
 Uy 482 DALVWVIAKSGIISOOOSKRIANI 406
 Db 856 IGRKQGVSSORFVQIETL 877

RESULT 49
 US 08 446 4088 6
 : Sequence 6: Application US/084464088
 : Patent No. 5724329
 : GENERAL INFORMATION:
 : APPLICANT: EVANS, RONALD W.
 : APPLICANT: MANGELSDORF, DAVID J.
 : TITLE OF INVENTION: RETINOID RECEPTOR COMPOSITIONS AND
 : TITLE OF INVENTION: METHODS
 : NUMBER OF SEQUENCES: 7
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: PRETTY, SCHROEDER, BROEGEMANN & CLARK
 : STREET: 444 SOUTH FLOWER STREET, SUITE 2000
 : CITY: LOS ANGELES
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 90071
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC DOS/MS DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/446,408B
 : FILING DATE: 08 NOV 1994
 : CLASSIFICATION: 445
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/944,454
 : FILING DATE: 21 AUG 1992
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: WO PCT/US91/00499
 : FILING DATE: 22 JAN 1991
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/478,071
 : FILING DATE: 09 FEB 1990
 : ATTORNEY/AGENT INFORMATION:
 : NAME: KETTER, STEPHEN E.
 : REGISTRATION NUMBER: 41,192
 : REFERENCE/WORK NUMBER: 941 9851
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 619 546 1995
 : TELEFAX: 619 546 9492
 : INFORMATION FOR SEQ ID NO: 6:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 463 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US 08 446 4088 6

Query Match 17.68; Score 444; Db 1; Length 463;
 Best Local Similarity 27.88; Prod. No. 2,40,49;
 Matches 145; Conserved Eve 69; Mismatches 205; Indels 76; Gaps 14;
 Db 4 NYSNINLLEGGGROTSNVLMP - - - - - PNRRETIKRRVSGNCASIV
 Uy 50 QKSWTFARSLNTL - - - - - PNRRETIKRRVSGNCASIV
 Db 64 LNLALSYRYVTSAMNPSPALAAPGILNLANPSSQLNVVSSSSDIKILRLGRLG 122
 Uy 65 TGGSKQALHFAVSLVYASGVYGVWSVKKKAFKRS1QGHINDYFVATNGG 148
 Db 124 NMVPSYSGS-LVHLFAVCPDSSSKHYGVYVCPCKCKPFRKELLYTQDNGG 181

Uy 149 NNSPSSGATGPGYFVWVNGSPRPGYVGVPPGSAFPGVWAKKPSNGH 198
 Db 182 LTPPGPNCVGVKPGYLVWVGNKRFVGP - - - - - GSPRPAFPA 241
 Uy 199 APEVRELLALSPDVLTLLEAPPVVLSRPSAPTEASMMMS1KLAKEVMS 258
 Db 242 EMPVERLEA - - - - - ELAVEKRFESYGMVENSTNDVNTDHAUKOLFIVE 282
 Uy 259 WAKK1GVVLESLFQVRLSPVWVYVMM1LMSSTHPPK1FAPV1LDEKQVE 318
 Db 283 WAKR1PESD1TEEDV1LKRMMNELLASFSHRSVVOG1LATV1LHVRSSANS-R 341
 Uy 419 GLETFPMLA-TTSPELEK1JREYLVKAM1LNSMVPV1VATV1A1SSPKRI 377
 Db 442 GVSSTFPAVITELVSKRMKQMKSE1G1A1V1FNDVAKG1NS1SEVTLREKVA1L 401
 Uy 478 NAVTDALVWVIAKSGIISOOOSKRIANIMLSHVRHANSKQEM1LMMK1KNV1YVDL 447
 Db 402 EAVTKO - - - - - KPEQNGRPAK1LRLPALRS1GLK1LH1FPEK1L1G1P1D1SF 451
 Uy 448 LEML 442
 Db 452 LEML 456

RESULT 40
 PCT US91 00399-6
 : Sequence 6: Application PCT/US9100399
 : GENERAL INFORMATION:
 : APPLICANT: Mangelsdorf, Dr., David J.
 : APPLICANT: EVANS Dr., Ronald W.
 : TITLE OF INVENTION: RETINOID RECEPTOR COMPOSITIONS AND
 : NUMBER OF SEQUENCES: 7
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: FITCH, EVEN, TAHIN & PLANNERY
 : STREET: 145 South LaSalle Street, Suite 900
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: USA
 : ZIP: 60603
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC DOS/MS DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US91/00399
 : FILING DATE: 19910122
 : CLASSIFICATION: 445
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/478,071
 : FILING DATE: 09 FEB 1990
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ketter, Mr., Stephen E.
 : REGISTRATION NUMBER: 41192
 : REFERENCE/WORK NUMBER: 50852
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (619) 552-1311
 : TELEFAX: (619) 552-0095
 : TELEX: 20 6566 PATLAW CRO
 : INFORMATION FOR SEQ ID NO: 6:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 463 amino acids
 : TYPE: AMINO ACID
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : PCT-US91-00399-6

Query Match 17.68; Score 444; Db 5; Length 463;
 Best Local Similarity 27.88; Prod. No. 2,40,49;
 Matches 145; Conserved Eve 69; Mismatches 205; Indels 76; Gaps 14;

```
0Y 2 NYS-----LPSNVTLFGPGKQOTSPVNLMT-----PGHL-SPLVWHQISSLHAAEP 49
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 4 NYSIFMKPFPTGPGSPGHTOSTMSPSVALPTCKMDSDHPSTDTTPVASAPRTLSAV-GTF 62
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 50 QKSPWFPAPQPHIT-----PVPETLKKVSGMRGASIV-----84
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 6 LNALSPYRVLTSMGPPSPALAAPGTINLVAPPSSQLNVVNSVSSSEPTKPLPCLPCLG 122
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 85 -----TGPGSKRIAHFGAAGSDYAGYHGYWMSCEGCAFFKRSIQCHNDYICPATNO* 138
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 123 MNMFSTSPSS LKRIHATGGRSSGRHGVVSTGRKGFRKLPKHLVYGRNEDW* 181
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 139 TIDNPKSGTACPIPGYGVVGVWGSGPPPGYPIVPPPGACDTHGACVAKRSQGH 198
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 182 LIEKROHRCQYCYVQKTIWGMKREAVQHR-----QKSHKASHPACA-----SSSH 231
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 199 AVKVELLLDAISPEQIVLTLLEAPPHVLISSPSAFTEASMMMSLTKIADKEIVHMTS 258
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 232 EDMVERILEA-----ELAVEPKTESYGDMMVENSTNDPVNICHADKQLPTLVE 282
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 259 WAKKIDGVELSLTQVRIIFSCMMFVIMMGIWKRSDHPRGKIIPAVHVLNDRGKVE 318
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 283 WAKRIPIHSDTLIEDVILLRAGNNELLASFHSKRSVVOGILLALGLVHRSAS*-K 341
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 319 GILETFMILA-TSKRETELQHKKEYLVKAMILLNSMYPIVTAFOHADSRLAHLL 377
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 342 GVGSITPDRLTELVSRRKMDQMSKSLGCLRAIVLFPDAKGLSNISVEVTLREKYATL 401
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 378 NAYTDALVWYIARSGTSSQSSMPANILMLLSHVPHASNGMHHLLNKKCKNVVPVYD 437
    |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 402 FAYTRQ -----KYPQPGPFARITIPALPSTGLKLEHLFFEXLIGDTIPIDSF 451
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0Y 438 LLEM 442
    |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
0b 452 LMEM 456
```

Search completed: June 16, 2003, 13:39:04
Job time : 24.4645 secs

•

•

•

43 352.5 14.0 746 9 US-09 042-488B-7 Sequence 7, Appl 1
 94 351.5 13.9 625 10 US-09 965-704-16 Sequence 16, Appl 1
 95 351.5 13.9 764 9 US-10 400-757-20 Sequence 20, Appl 1
 96 349.5 13.9 395 9 US-10 151-542A-21 Sequence 21, Appl 1
 97 349.5 13.7 446 12 US-10 614-B2-2 Sequence 2, Appl 1
 98 344.5 13.7 446 10 US-09 909-446-2 Sequence 2, Appl 1
 99 344.5 13.7 446 10 US-09 909-425-2 Sequence 2, Appl 1
 100 344.5 13.6 446 10 US-09 909-426-2 Sequence 2, Appl 1
 101 341.5 13.6 746 9 US-10 400-757-18 Sequence 18, Appl 1
 102 340 13.5 514 10 US-09 965-704-59 Sequence 59, Appl 1
 103 339.5 13.5 746 9 US-09 042-488B-5 Sequence 5, Appl 1
 104 338.5 13.4 1041 9 US-09 042-488B-9 Sequence 9, Appl 1
 105 333 13.2 61 9 US-09 814-604-21 Sequence 21, Appl 1
 106 333 13.2 61 9 US-09 922-226-7 Sequence 7, Appl 1
 107 333 13.2 61 9 US-09 922-226-7 Sequence 15, Appl 1
 108 330 13.1 367 9 US-10 151-542A-2 Sequence 2, Appl 1
 109 328 13.0 546 9 US-10 295-470-2 Sequence 2, Appl 1
 110 324.5 12.8 440 10 US-09 965-704-71 Sequence 71, Appl 1
 111 317.5 12.6 412 10 US-09 965-704-11 Sequence 11, Appl 1
 112 317.5 12.6 412 10 US-09 965-704-12 Sequence 12, Appl 1
 113 315 12.5 60 10 US-09 864-761-4015 Sequence 4015, A
 114 302.5 12.0 69 10 US-09 952-559-3 Sequence 3, Appl 1
 115 300 11.9 476 9 US-10 188-721-1 Sequence 1, Appl 1
 116 298 11.8 448 10 US-09 760-464-1 Sequence 1, Appl 1
 117 298 11.8 469 10 US-10 155-479-2 Sequence 2, Appl 1
 118 297 11.8 359 9 US-09 760-464-9 Sequence 9, Appl 1
 119 292.5 11.6 477 9 US-10 182-232-3 Sequence 3, Appl 1
 120 290 11.5 264 9 US-10 043-487-254 Sequence 254, Appl 1

ALIGNMENTS

RESULT 1
 US-10 198-785-2
 Sequence 2, Application US/10198785
 Publication No. US2004002224A1
 GENERAL INFORMATION:
 APPLICANT: OLYMPUS OPTICAL CO., LTD.
 TITLE OF INVENTION: Method of detecting binding reaction between protein and host sub
 FILE REFERENCE: 781 0280548
 CURRENT APPLICATION NUMBER: US/10/198-785
 PRIOR FILING DATE: 2002 07 19
 PRIOR APPLICATION NUMBER: JP2001-229444
 PRIOR FILING DATE: 2001 07 19
 PRIOR APPLICATION NUMBER: JP2001-221962
 PRIOR FILING DATE: 2001 07 24
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: Patent In Vei., 2.0
 SEQ ID NO 2
 LENGTH: 540
 TYPE: PRT
 ORGANISM: Homo sapiens
 US 10 198 785 2

Query Match 100.0% Score: 2522, 100.0, Length: 540
 Post Local Similarity 100.0% Prod. No. 10-2162
 Matches: 4772 Conserved: 02 Mismatches: 92 Indels: 0 Gaps: 92
 1 MNY:ETPVNLTDEGKGTSTNVLTGHSPLVYHQLSHUYAEKSKWFEASL 60
 114 MNY:ETPVNLTDEGKGTSTNVLTGHSPLVYHQLSHUYAEKSKWFEASL 113
 61 EHTLVKRETLKRVSNRGASVTVGSKRQAMPVAVSTVASTYHYVWSSKAEF 120
 114 EHTLVKRETLKRVSNRGASVTVGSKRQAMPVAVSTVASTYHYVWSSKAEF 173
 121 KRSYQGNHYTCATNGCTTKRKRKSVAVTLKRYLWYQWNVSSPSPKRYKVKR 300
 174 KRSYQGNHYTCATNGCTTKRKRKSVAVTLKRYLWYQWNVSSPSPKRYKVKR 253
 181 EADYQGNHYTCATNGCTTKRKRKSVAVTLKRYLWYQWNVSSPSPKRYKVKR 240

106 244 SAEQLEWAKAKRSQSHARVPELLEALDEQVLTLEAFYVLTSSPAETLAS 293
 241 MMYSTELAKELVIMISAKKIDPEVELSLPEQVLLKSCWVEYIMMLKRSIDHKK 300
 294 MMYSTELAKELVIMISAKKIDPEVELSLPEQVLLKSCWVEYIMMLKRSIDHKK 353
 401 LLEAFQVLDRECKVEGTELEFQMLATTSFRELKQHEEYLVKAMLLNSMYEL 360
 454 LLEAFQVLDRECKVEGTELEFQMLATTSFRELKQHEEYLVKAMLLNSMYEL 413
 401 VIAVQVDSKRLAHNAVVALVYVAKSGTSSQGSRLANLMLSHKVASNRM 420
 414 VIAVQVDSKRLAHNAVVALVYVAKSGTSSQGSRLANLMLSHKVASNRM 473

RESULT 2

US-09-854-033-2
 Patent No. US20020100068A1
 GENERAL INFORMATION:
 APPLICANT: CHAMBERLAIN, PIERRE
 TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
 FILE REFERENCE: 065691/0222
 CURRENT APPLICATION NUMBER: US/09/854,033
 PRIOR FILING DATE: 2001 05 11
 PRIOR APPLICATION NUMBER: FR 96/712570
 PRIOR FILING DATE: 2000-10-03
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: Patent In Vei., 2.1
 SEQ ID NO 2
 LENGTH: 595
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-854-033-2

Query Match 47.7% Score: 1204, 100.0, Length: 595
 Post Local Similarity 50.9% Prod. No. 66-992
 Matches: 216 Conserved: 82 Mismatches: 142 Indels: 54 Gaps: 102
 10 TLESGPGRVSTSTNVLTGHSPLVYHQLSHUYAEKSKWFEASL 50
 76 TLESGPGRVSTSTNVLTGHSPLVYHQLSHUYAEKSKWFEASL 135
 51 KSTWVEEAKSLHTPLVYHQLSHUYAEKSKWFEASL 103
 136 KSTWVEEAKSLHTPLVYHQLSHUYAEKSKWFEASL 192
 104 SGHYGVWSTDEKKAERKRSYQGNHYTCATNGCTTKRKRKSVAVTLKRYLWYQWNV 163
 193 SGHYGVWSTDEKKAERKRSYQGNHYTCATNGCTTKRKRKSVAVTLKRYLWYQWNV 252
 164 GSNRRRGVYLVKRSQSHARVPELLEALDEQVLTLEAFYVLTSSPAETLAS 207
 253 GSNRRRGVYLVKRSQSHARVPELLEALDEQVLTLEAFYVLTSSPAETLAS 307
 209 EATSPEDVYLTLEAFYVLTSSPAETLAS 265
 408 EATSPEDVYLTLEAFYVLTSSPAETLAS 466
 206 EVELSLPEQVLLKSCWVEYIMMLKRSIDHKK 325
 407 EVELSLPEQVLLKSCWVEYIMMLKRSIDHKK 466
 407 EVELSLPEQVLLKSCWVEYIMMLKRSIDHKK 466
 427 EVELSLPEQVLLKSCWVEYIMMLKRSIDHKK 486

385 WVAIAKSGISGGQSMRIANIIMLSHVHASKKMEHLLNKKKNVVPYDILLEMNA 444
487 IHMAKAGGTLIGQOHQAGLILSHIRHMSKMEHLYSMKKNVVPELIDLEMLIA 546
445 HVL 447
547 HRL 549

RESULT 4
US-10-052-092-40
Sequence 30, Application US/10052092
Publication No. US2003002779A1
GENERAL INFORMATION:
APPLICANT: Fougna, Suzanne
APPLICANT: Alfred, D.
APPLICANT: Hopp, Torsten A.
APPLICANT: O'Connell, Peter
TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapies
FILE REFERENCE: P02102052
CURRENT APPLICATION NUMBER: US/79/052,092
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: us 60/242,990
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/104,018
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patent In version 3.1
SEQ ID NO 30
LENGTH: 595
TYPE: PRT
ORGANISM: hamster
US-10-052-092-40

Query Match 47.7% Score 1297 DB 9 Length 595
Best Local Similarity 50.7% Prod. No. 90-99
Matches 246 Conservative 87 Mismatches 122 Indels 30 Gaps 10

16 EGRGTTSTN VLMPTGCHSP LVVHRSLSHVAPKSPKFAFSLFHTLVNPFET 71
97 FOLNSVSPSLMLHPPPOLSTFLPHRLQVPTYLENERSAIA VFGSPFAFVSNG 156
72 KKKVGNRCASPVTPGSK RDAHFVAVSDVASYHGVWSGFCGKAFKFKSTG 126
156 NKKSGHFPRISSSFKSGMAMFVKETPYCAVNPVASYHYHGVWSGFCGKAFKFKSTG 215
127 HNYTPRATNGCTTKRKKRKSQVTPRPVYVGVWKKVGSPPRPVYVYVFGPSATF 194
216 HIRHMGATVAGCTTKRKKRKSQVTPRPVYVGVWKKVGSPPRPVYVYVFGPSATF 215
185 --GIFACAKRPSG--HAPVPELLEDAISPEQVLTLEEDVHVLIS-ET 272
276 GRNMGSGMGRATNIWPSPLVLRKTKKNSPAL--SLTADQVMSALLMEPLITYEYH 333
233 SAFTTASMMGLTKLAKELVHMSAKKIPGFVLSLQVRLLESQWMLVLMGIMM 292
434 SRPSEASMMGLTLNLADEFLVHIMAKKVFSGFQNIHQVHLIFVAMI FLMGLIM 493
293 KSTHPRKLTFAPIVTPRPGKVFQITFIHMLLATSPEFLKLGREFTLVYAMLI 452
494 KSMHPRKLTFAPIVTPRPGKVFQITFIHMLLATSPEFLKLGREFTLVYAMLI 453
453 LNSMPLVAT--QDAUSPKRIANIINAVTALWVAIAKSGISGGQSMRIANIIMLSH 411
454 LNSGVYTFELSTLKLSEFKDILHVLIDKLTDTPLIHLMAKAGLTLQVGHRLAQLILSH 513
412 VRAASNGMEHLLNKKKNVVPYDILLEMNAHVI PRKKSITTSNGSPATFVSNG 471
514 IRHMSNGMEHLLNKKKNVVPYDILLEMNAHRLH--TPVSRMGVSPEDPSQTLT 570
472 QNPOS 476
571 INSTS 575

385 WVAIAKSGISGGQSMRIANIIMLSHVHASKKMEHLLNKKKNVVPYDILLEMNA 444
487 IHMAKAGGTLIGQOHQAGLILSHIRHMSKMEHLYSMKKNVVPELIDLEMLIA 546
445 HVL 447
547 HRL 549

RESULT 4
US-10-096-710-1
Sequence 1, Application US/10096710
Patent No. US20020164581A1
GENERAL INFORMATION:
APPLICANT: Chang, Chawshuan
TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS THAT
TITLE OF INVENTION: DO NOT ATTENUATE THE PROTECTIVE EFFECTS OF BSKRGEN
FILE REFERENCE: 21108,601002
CURRENT APPLICATION NUMBER: US/10/096,710
PRIOR FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/275,229
PRIOR FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 595
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Is a 5' piece of Artificial Sequence/Note -
US-10-096-710-1

Query Match 47.4% Score 1195 DB 9 Length 595
Best Local Similarity 50.7% Prod. No. 80-98
Matches 245 Conservative 81 Mismatches 103 Indels 54 Gaps 10

10 TNEGGGPGP-----TTSPPVLMPTGCHSP LVVHRSLSHVAPKSPKFAFSLFHTLVNPFET 50
76 TGLPYGGSFAAFAGSMGLGFPPLNSVSPSLMLHPPPOLSTFLPHRLQVPTYLENERSAIA 135
51 KSIW--FAKSIHTITVNETIKRVSNG--ASVYLG-- SKPLAH--CAV--SYA 103
136 ESYTVEAFSHVAPVY--NSINHGQSGPPLASTINHGSMAMSAKELVYCAVNPVASY 192
104 SCYHGVWSGFCGKAFKFKSTGNIYVTAIAGCTTIIRPKRSQVATPLPCTFVGWKK 163
193 SHVHGVWSGFCGKAFKFKSTGNIYVTAIAGCTTIIRPKRSQVATPLPCTFVGWKK 252
164 GNSPFRVYVLPVPSADPQRIATKAKFSTHAPVY-----ETIT 207
253 GQPKKQDQDQPMILKRWQDRI--QDQGVWSGCMQPMALWQDQMTKPSYKSLA 307
208 DALSTEDVLLLEFATVHVLISR--SAFTTASMMGLTKLAKELVHMSAKKIPGFVLSLQVRLLESQWMLVLMGIMM 265
434 SRPSEASMMGLTLNLADEFLVHIMAKKVFSGFQNIHQVHLIFVAMI FLMGLIM 493
293 KSTHPRKLTFAPIVTPRPGKVFQITFIHMLLATSPEFLKLGREFTLVYAMLI 452
494 KSMHPRKLTFAPIVTPRPGKVFQITFIHMLLATSPEFLKLGREFTLVYAMLI 453
453 LNSMPLVAT--QDAUSPKRIANIINAVTALWVAIAKSGISGGQSMRIANIIMLSH 411
454 LNSGVYTFELSTLKLSEFKDILHVLIDKLTDTPLIHLMAKAGLTLQVGHRLAQLILSH 513
412 VRAASNGMEHLLNKKKNVVPYDILLEMNAHVI PRKKSITTSNGSPATFVSNG 471
514 IRHMSNGMEHLLNKKKNVVPYDILLEMNAHRLH--TPVSRMGVSPEDPSQTLT 570
472 QNPOS 476
571 INSTS 575


```

1b 193 SCYHYGVWSFEGKAFKFKRSJGHNQYVATNCTIDKMKRSQVAKRLKRYEYVGMK 252
UY 164 GCSRRERGVRLVBRORSADPOLHFCAGKAKRSQSHAPVR ----- ELL 207
DB 254 GGIKKRKRGGKRLKIRKRDND --- GGRKRGVSAAGMKRANMLWSTLMKRSKKNSTA 407
1b 254 GGIKKRKRGGKRLKIRKRDND --- GGRKRGVSAAGMKRANMLWSTLMKRSKKNSTA 407
UY 208 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK193 245
DB 408 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK193 245
1b 408 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK193 245
UY 266 FVELSLFDQVRLLESQVWMEVLMGIMKRSIDHCKLFAVDLYLDKRGKVGSLLEFD 425
DB 467 FVELSLFDQVRLLESQVWMEVLMGIMKRSIDHCKLFAVDLYLDKRGKVGSLLEFD 426
1b 467 FVELSLFDQVRLLESQVWMEVLMGIMKRSIDHCKLFAVDLYLDKRGKVGSLLEFD 426
UY 426 MLATTSRPRELKIQRKRYLYVAKMILNNSMYPLVAT-QDADSSRKLAHLNAVTDAL 484
DB 427 MLATTSRPRELKIQRKRYLYVAKMILNNSMYPLVAT-QDADSSRKLAHLNAVTDAL 486
1b 427 MLATTSRPRELKIQRKRYLYVAKMILNNSMYPLVAT-QDADSSRKLAHLNAVTDAL 486
UY 485 VVVIKSGTSSQOOSMRANLMLSLVHBSKMGHEHLMKCKNVVYVILLLEMDNA 444
DB 487 ILMKAKGTLTQOQHRLAQLLLLSHRMSKMGHEHLYSMCKKNVYVILLLEMDNA 546
1b 487 ILMKAKGTLTQOQHRLAQLLLLSHRMSKMGHEHLYSMCKKNVYVILLLEMDNA 546
UY 445 HVL 447
DB 547 HVL 549

```

```

RESULT 10
US 09 943 267A 2
Sequence 2, Application US/09943267A
Patient No. US20020124095A1
GENERAL INFORMATION:
APPLICANT: Kalush, Francis et al.
TITLE OF INVENTION: Estrogen receptor alpha variants and
FILE REFERENCE: C1000258C14
CURRENT APPLICATION NUMBER: US/09/943,267A
PRIOR APPLICATION NUMBER: 2001-04-05
PRIOR FILING DATE: 1999-10-20
PRIOR APPLICATION NUMBER: 60/184756
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: 09/692414
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/768184
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 09/804076
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/826314
PRIOR FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 595
TYPE: PRT
ORGANISM: HUMAN
US 09 943 267A 2

```

```

Query Match 47.4%; Score 1195; DB 10; Length 595;
Host Local Similarity 50.7%; Prod. No. 3,86-98;
Matches 249; Conservative 81; Mismatches 103; Indels 54; Gaps 10;
1b 10 INLEGGRO -----TTSNVLMTFCHLSP-LVHQLSHYAEQ 50
DB 76 TGLAYDQSEAAVNSMTCGPPQNSVSPSLVLIHPPQVSPFLHBRQGVVYVLENE 145
UY 51 KSWP -EAKSEHTLPVNRRTKKRVSNRCASPTVPG -SKRAHFAVASYVA 104
DB 136 PSYTVFAVAPVAYFP -NSDRKRGGRKRLASTIDKMSAMESAKETRYCAVNDVA 192
UY 104 SCYHYGVWSFEGKAFKFKRSJGHNQYVATNCTIDKMKRSQVAKRLKRYEYVGMK 164
DB 193 SCYHYGVWSFEGKAFKFKRSJGHNQYVATNCTIDKMKRSQVAKRLKRYEYVGMK 252

```

```

UY 164 GCSRRERGVRLVBRORSADPOLHFCAGKAKRSQSHAPVR ----- ELL 207
DB 254 GGIKKRKRGGKRLKIRKRDND --- GGRKRGVSAAGMKRANMLWSTLMKRSKKNSTA 407
UY 208 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK193 245
DB 408 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK193 245
1b 408 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK193 245
UY 266 FVELSLFDQVRLLESQVWMEVLMGIMKRSIDHCKLFAVDLYLDKRGKVGSLLEFD 425
DB 467 FVELSLFDQVRLLESQVWMEVLMGIMKRSIDHCKLFAVDLYLDKRGKVGSLLEFD 426
1b 467 FVELSLFDQVRLLESQVWMEVLMGIMKRSIDHCKLFAVDLYLDKRGKVGSLLEFD 426
UY 426 MLATTSRPRELKIQRKRYLYVAKMILNNSMYPLVAT-QDADSSRKLAHLNAVTDAL 484
DB 427 MLATTSRPRELKIQRKRYLYVAKMILNNSMYPLVAT-QDADSSRKLAHLNAVTDAL 486
1b 427 MLATTSRPRELKIQRKRYLYVAKMILNNSMYPLVAT-QDADSSRKLAHLNAVTDAL 486
UY 485 VVVIKSGTSSQOOSMRANLMLSLVHBSKMGHEHLMKCKNVVYVILLLEMDNA 444
DB 487 ILMKAKGTLTQOQHRLAQLLLLSHRMSKMGHEHLYSMCKKNVYVILLLEMDNA 546
1b 487 ILMKAKGTLTQOQHRLAQLLLLSHRMSKMGHEHLYSMCKKNVYVILLLEMDNA 546
UY 445 HVL 447
DB 547 HVL 549

```

```

RESULT 11
US 10 052 092-31
Sequence 31, Application US/10052092
Patient No. US20030027778A1
GENERAL INFORMATION:
APPLICANT: Fuqua, Suzanne
APPLICANT: Allrod, D.
APPLICANT: Hopp, Kristen A.
TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapy
FILE REFERENCE: P021020S2
CURRENT APPLICATION NUMBER: US/10/052,092
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/262,990
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/404,018
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 49
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 41
LENGTH: 595
TYPE: PRT
ORGANISM: PIG
US 10 052 092-31

```

```

Query Match 47.3%; Score 1194; DB 9; Length 595;
Host Local Similarity 49.1%; Prod. No. 5,76-98;
Matches 249; Conservative 86; Mismatches 110; Indels 62; Gaps 14;
1b 10 INLEGGRO -----TTSNVLMTFCHLSP-LVHQLSHYAEQ 50
DB 76 TGLAYDQSEAAVNSMTCGPPQNSVSPSLVLIHPPQVSPFLHBRQGVVYVLENE 145
UY 51 KSWP -EAKSEHTLPVNRRTKKRVSNRCASPTVPG -SKRAHFAVASYVA 104
DB 136 PSYTVFAVAPVAYFP -NSDRKRGGRKRLASTIDKMSAMESAKETRYCAVNDVA 192
UY 104 SCYHYGVWSFEGKAFKFKRSJGHNQYVATNCTIDKMKRSQVAKRLKRYEYVGMK 164
DB 193 SCYHYGVWSFEGKAFKFKRSJGHNQYVATNCTIDKMKRSQVAKRLKRYEYVGMK 252
1b 164 GCSRRERGVRLVBRORSADPOLHFCAGKAKRSQSHAPVR ----- ELL 207
UY 208 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK1 264
DB 254 GGIKKRKRGGKRLKIRKRDND --- GGRKRGVSAAGMKRANMLWSTLMKRSKKNSTA 407
1b 254 GGIKKRKRGGKRLKIRKRDND --- GGRKRGVSAAGMKRANMLWSTLMKRSKKNSTA 407
UY 408 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK1 264
DB 408 IATSPGQVLTLEAPPHVILSR -PSAPFTASMMNSTKLADREYVHMISWARK1 264

```

```

408 1--SITADQMISALFAPPP--11YSHYDPTREISASMMCLITNATAPVIMIMAKPV 364
QY 264 PGEVELSLFDVLLFSCYMEVIMMGIIMKSTIDHPSKIFAPRIVLELEEEVEGILET 423
DB 365 PGEVLDSLDVLLFSCYMEVIMMGIIMKSTIDHPSKIFAPRIVLELEEEVEGILET 424
QY 324 FOMLATTSFRELKQREYLCVKAMILLNSMYLVLAIVQADSSKRLANLINVVD 482
DB 425 FOMLATTSFRELKQREYLCVKAMILLNSMYLVLAIVQADSSKRLANLINVVD 484
QY 483 AIVWIAKSTISQVQSMPIANIMLSHVPHASNKMEHLLMCKGVVAVYDILLML 442
DB 485 ILIHMAKASTIGQPHPLAQLILLSSPHNSKGMHLYNMCKGVVAVYDILLML 544
QY 443 NAHVLKCKKSTIGSECSAEDSKSE 469
DB 545 DAHRLH-APTINDG---PEPEDMSQSO 567

RESULT 12
US-09-893-666A-2
Sequence 2, Application US/09893666A
Patent No. US20020129486A1
GENERAL INFORMATION:
APPLICANT: YAMASHITA, ICHIRO
TITLE OF INVENTION: High estrogen sensitive medaka fish
FILE REFERENCE: 21021728-420-7249 0
CURRENT APPLICATION NUMBER: US/09/893, 666A
CURRENT FILING DATE: 2002-02-12
PRIOR APPLICATION NUMBER: JP 2000 247729
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent in version 3.1
SEQ ID NO 2
LENGTH: 575
TYPE: PRI
ORGANISM: Oryzias latipes
US-09-893-666A-2

Query Match 47.3%; Score 1192; DB 10; Length 575;
Best Local Similarity 49.5%; Pred. No. 6,7e 98;
Matches 260; Conservative 74; Mismatches 121; Indels 70; Gaps 14;

QY 3 YSLP-----SNVNLGGPGROITSNVLPGLSPILVPHQLSHLYABQSP 53
DB 43 YSAPLETNPPSGSLSQSGP-----TSPLVVPSSRPLSPFM--HPSHHTLETTSP 96
QY 54 WCFARSLHPTLVNRETLLKRVSGNR---CASPVIGPGS-----KRIAPHCAVCSQYA 103
DB 97 VYRSS-----HQSASPELQSGSPEDTSLGAGAGAGGFEMAKTFEFAVSHYA 148
QY 104 SCGHVYVWSPCKKAFKPSIGGHNVYTPPATNGTTPNPPSCQVPLPQVYVWY 163
DB 149 SCGHVYVWSPCKKAFKPSIGGHNVYTPPATNGTTPNPPSCQVPLPQVYVWY 208
QY 164 GQSHREKQYRLVRQK-----SADPOLFCAGKAKRS--GQHAPRVHLL 207
DB 209 GQVAKKQ--IRHLPRIKHPITGVGDCIKVYKQGEIKTVHITGKRBSSTGGGGRSSGLSLV 256
QY 208 DALSPGCVITTLFATPVATSEF--SAPTFASMMKSTIKALVHIVAM--WAP--TFC 265
DB 267 TSLPPECVILLQVAPP--TICSPQKLSPEYTEVIMTLLTSMARKEIVHIMAMAKLEQ 425
QY 266 FVELSLFDVLLFSCYMEVIMMGIIMKSTIDHPSKIFAPRIVLELEEEVEGILET 425
DB 426 FVELSLFDVLLFSCYMEVIMMGIIMKSTIDHPSKIFAPRIVLELEEEVEGILET 485
QY 424 MIACTSPSPFPIQHRYVYVAM--LNSMYP--VAVAG--MSSSLFALL--NAVLA-- 434
DB 386 MLATASRPVILKLEEFVCKATILHNSGASPTCTIMEPLHNSAVQSMPTITVAL 445
QY 385 VWVIAKSTISQVQSMPIANIMLSHVPHASNKMEHLLMCKGVVAVYDILLML 444

```

```

446 1HLSQSYLAAGVAGGAGLLLSLHLSHNSKMKHLYSMCKRVNVTYLLMLLA 505
QY 445 EVL-----KCKS-----SLSGDSYALSKSGSGNP 474
DB 506 HRLHVPFVAGSLQVWHPSPSSSGSGGIALSISASGRLESP 560

RESULT 14
US-09-903-876-2
Sequence 2, Application US/09903876
Patent No. US2002022587A1
GENERAL INFORMATION:
APPLICANT: Xu, Zhang-Bao
APPLICANT: Somers, William
APPLICANT: Akopian, Talos
APPLICANT: Hsiao, Chu-Lai
APPLICANT: Umalla, Raymond
TITLE OF INVENTION: CRYSTAL STRUCTURE OF ESTROGEN RECEPTOR-B COMPLEX
FILE REFERENCE: 97174/9
CURRENT APPLICATION NUMBER: US/09/903, 876
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: US 60/217, 834
PRIOR FILING DATE: 2000-07-12
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent in version 3.0
SEQ ID NO 2
LENGTH: 239
TYPE: PRI
ORGANISM: Homo sapiens
US-09-903-876-2

Query Match 47.2%; Score 1190; DB 10; Length 239;
Best Local Similarity 99.2%; Pred. No. 2,9e 98;
Matches 239; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 407 LMAISPOLVLLLAIFPRVILSKSAVLEASMMSLKIALDEELVMSMAKIN 466
DB 1 LMAISPOLVLLLAIFPRVILSKSAVLEASMMSLKIALDEELVMSMAKIN 60
QY 267 VELSLFDVLLFSCYMEVIMMGIIMKSTIDHPSKIFAPRIVLELEEEVEGILET 326
DB 61 VELSLFDVLLFSCYMEVIMMGIIMKSTIDHPSKIFAPRIVLELEEEVEGILET 120
QY 327 TATSPSPFPIQHRYVYVAM--LNSMYPVATGDAISSPKIALMLNAVDAVW 386
DB 121 TATSPSPFPIQHRYVYVAM--LNSMYPVATGDAISSPKIALMLNAVDAVW 179
QY 387 VIAKSTISQVQSMPIANIMLSHVPHASNKMEHLLMCKGVVAVYDILLML 446
DB 180 VIAKSTISQVQSMPIANIMLSHVPHASNKMEHLLMCKGVVAVYDILLML 444
QY 447 L 447
DB 239 L 239

RESULT 14
US-10-052-092-12
Sequence 12, Application US/10052092
Publication No. US2003002778A1
GENERAL INFORMATION:
APPLICANT: Futaba, Suzanne
APPLICANT: Allred, D.
APPLICANT: Hopp, Torsten A.
APPLICANT: O'Connell, Peter
TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapy
FILE REFERENCE: P02102082
CURRENT APPLICATION NUMBER: US/10/052, 092
CURRENT FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/262, 990
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/304, 018

```

PRIOR FILING DATE: 2001 07 09
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: Patent In version 4.1
 SEQ ID NO: 12
 LENGTH: 701
 TYPE: PRT
 ORGANISM: human
 US 10 052 092 12

Query Match 45.4% Score 1145 DB 9 Length 701

Host Local Similarity 42.4% Pred. No. 146-93
 Matches 250 Conservative 80 Mismatches 99 Indels 160 Gaps 12

```

QY 10 TLEKGGGCGG-----TSPVAVMTPECHLSP-LVNHQSLSHYAEQ 50
   1 111 1 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 76 TGLPYGGSEAAAGSMGLGGPPLNSVSPMLLHPPLQSPPLQPGGGVYVLENE 135
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 51 KSPW--EASLPTLVNRETKRKYSCKKASPTGQ-----SKDAHPAVGSYA 103
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 136 PSYTVFAAGPAPYK-----NSDKRGQDREPLASINRKSMAKESAKETRYAVNYVA 192
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 104 SGHYGVWSTGECAPKPKRSLOCHNDYICPATNQTLIDKRRKSVQACMLKRYGVGVK 163
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 193 SGHYGVWSTGECAPKPKRSLOCHNDYICPATNQTLIDKRRKSVQACMLKRYGVGVK 252
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 164 GSPPEPGYPLVPGPSAIDPQIICACAKPSCHAPV-----ELLL 207
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 253 GSPPEPGYPLVPGPSAIDPQIICACAKPSCHAPV-----ELLL 207
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 208 DAESEGLVLTLEAFTHVLSR--ESAFETLASMMSLTELADRELYHMSWAKETG 265
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 408 LSTADQVNSALIDAEFT--LTSYDRETRPSASMMGLTINLADRELYHMIWAKETG 366
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 266 FVELSTGCVGLDS--WMEVIMMLMKSLDREKLEAFQVLDHDEPKRYVETLETH 425
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 467 FVELSTGCVGLDS--WMEVIMMLMKSLDREKLEAFQVLDHDEPKRYVETLETH 426
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 426 MLATTSKPEELKIGKEEYLVKAMLLNSSMYEVIAT-----QD----- 466
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 427 MLATTSKPEELKIGKEEYLVKAMLLNSSMYEVIAT-----QD----- 466
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 367 -----ADSR----- 371
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 487 IHMAKATITLQCHQHLAQLLLSHIRHRNGCKVEGMVEIFDMLATSSRFMMNL 546
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 372 -----KLAHLNAVVDALWVYIAKSGISSQ 398
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 547 QEEFVLTSLTLLNSGVYTLSSLSKSEKHLIRVLDKLTDTLHMAKAGLTQ 606
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 499 SMRLANILMLSHVHANSKGMELNMRKNVVPYDILLEMNAHL 447
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 607 HQLAQLLLLSHIRHNSKGMELNMRKNVVPYDILLEMNAHL 655
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

RESULT 15

US 10 211 613 1
 Sequence 1, Application US/10211613
 Publication No. US200 600966A1A1
 GENERAL INFORMATION:
 APPLICANT: CELL SIGNALING TECHNOLOGY, INC.
 APPLICANT: SMITH, Bradley
 APPLICANT: GROSS, Kathleen
 APPLICANT: WU, Jiong
 TITLE OF INVENTION: Monoclonal Antibodies Specific for Phosphorylated Estrogen Receptor
 FILE REFERENCE: CST 202
 CURRENT APPLICATION NUMBER: US/10/211,613
 PRIOR FILING DATE: 2002-08-02
 PRIOR APPLICATION NUMBER: US 60/410,066
 PRIOR FILING DATE: 2001-08-03
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: Patent In version 4.1
 SEQ ID NO: 1

LENGTH: 701
 TYPE: PRT
 ORGANISM: Homo sapiens
 US 10 211 613 1

Query Match 45.4% Score 1145 DB 9 Length 701

Host Local Similarity 42.4% Pred. No. 146-93
 Matches 250 Conservative 80 Mismatches 99 Indels 160 Gaps 12

```

QY 10 TLEKGGGCGG-----TSPVAVMTPECHLSP-LVNHQSLSHYAEQ 50
   1 111 1 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 76 TGLPYGGSEAAAGSMGLGGPPLNSVSPMLLHPPLQSPPLQPGGGVYVLENE 135
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 51 KSPW--EASLPTLVNRETKRKYSCKKASPTGQ-----SKDAHPAVGSYA 103
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 136 PSYTVFAAGPAPYK-----NSDKRGQDREPLASINRKSMAKESAKETRYAVNYVA 192
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 104 SGHYGVWSTGECAPKPKRSLOCHNDYICPATNQTLIDKRRKSVQACMLKRYGVGVK 163
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 193 SGHYGVWSTGECAPKPKRSLOCHNDYICPATNQTLIDKRRKSVQACMLKRYGVGVK 252
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 164 GSPPEPGYPLVPGPSAIDPQIICACAKPSCHAPV-----ELLL 207
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 253 GSPPEPGYPLVPGPSAIDPQIICACAKPSCHAPV-----ELLL 207
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 208 DAESEGLVLTLEAFTHVLSR--ESAFETLASMMSLTELADRELYHMSWAKETG 265
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 408 LSTADQVNSALIDAEFT--LTSYDRETRPSASMMGLTINLADRELYHMIWAKETG 366
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 266 FVELSTGCVGLDS--WMEVIMMLMKSLDREKLEAFQVLDHDEPKRYVETLETH 425
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 467 FVELSTGCVGLDS--WMEVIMMLMKSLDREKLEAFQVLDHDEPKRYVETLETH 426
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 426 MLATTSKPEELKIGKEEYLVKAMLLNSSMYEVIAT-----QD----- 466
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 427 MLATTSKPEELKIGKEEYLVKAMLLNSSMYEVIAT-----QD----- 466
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 367 -----ADSR----- 371
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 487 IHMAKATITLQCHQHLAQLLLSHIRHRNGCKVEGMVEIFDMLATSSRFMMNL 546
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 372 -----KLAHLNAVVDALWVYIAKSGISSQ 398
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 547 QEEFVLTSLTLLNSGVYTLSSLSKSEKHLIRVLDKLTDTLHMAKAGLTQ 606
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 499 SMRLANILMLSHVHANSKGMELNMRKNVVPYDILLEMNAHL 447
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
DB 607 HQLAQLLLLSHIRHNSKGMELNMRKNVVPYDILLEMNAHL 655
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

RESULT 16
 US 10 259 864 4
 Sequence 4, Application US/10259864
 Publication No. US200 400764A1A1
 GENERAL INFORMATION:
 APPLICANT: Haver, Gordon L.
 TITLE OF INVENTION: Superfamily Receptor Chimeras, Translocation Assay for Support
 TITLE OF INVENTION: Receptor Ligands, and Methods and Kits for Detecting and Char
 FILE REFERENCE: 24799
 CURRENT APPLICATION NUMBER: US/10/259,864
 PRIOR APPLICATION NUMBER: US 60/425,178
 PRIOR FILING DATE: 2002-09-30
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patent In version 4.1
 SEQ ID NO: 4
 TYPE: PRT
 ORGANISM: Rat/human Chimera
 US 10 259 864 4

Query Match 48.0% Score 958 DB 5 DB 9 Length 1099


```

TYPE: PRI
ORGANISM: Homo sapiens
US-09-905-176-14

Query Match
Best Local Similarity 30.58; Score 768; DB 10; Length 241;
Pred No 1 4a-60;
Matches 144; Conservative 56; Mismatches 37; Indels 4; Gaps 4;

QY 210 LSPDQVLTLEAFPHVLSR--PSAPFTASMMSTLKALKEIVHIMSKAKKIPGV 267
DB 1 LADQVVSALLDAPP--IIVSEYDPTPEFSASMMGLTNLAPELVHIMNAKRVPGV 59
QY 268 FLSFPGVRIIFSGMMFVIMGIMWPSIDHP--KLIFAPPLVILPDGKVEGTEEPML 327
DB 60 DLTIDVHLLERAMLEFLIMGLVWRSMPKCLIFAPMLIDPNSGKVEGVEIFDML 119
QY 328 LATSPREPLKIQKREYICVKAMILLNSMYVIVAT--GDASSPKLAHLNAVTVLAV 386
DB 120 LATSPREPMNLQEEFVTLKSTILNSVYTFLSSTKSLPKKHIIHVLTKITDTLH 179
QY 387 VIAKSGISSQOOSKRLANLMLLSHYBANSKGMHELLNMKKKNVYVYDILLMLNAH 446
DB 180 LMAKATLIGQGHRLAQIILISHIRMSNKGMEHLVSMKKKNVYVYDILLMLNAH 289
QY 447 L 447
DB 240 L 240

RESULT 22
US-10-052-092-10
Sequence 10; Application US/10052092
Publication No. US2004002778A1
GENERAL INFORMATION:
APPLICANT: Fujita, Suzanne
APPLICANT: Allied, D.
APPLICANT: Hopp, Torsten A.
APPLICANT: O'Connell, Peter
TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapy
FILE REFERENCE: P02102982
CURRENT APPLICATION NUMBER: US/10/052,092
CURRENT FILING DATE: 2002-01-19
PRIOR APPLICATION NUMBER: US 60/262,990
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/304,018
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patent version 3.1
SEQ ID NO 10
LENGTH: 547
TYPE: PRI
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Synthetic Peptide Encoded by the Coding Region
US-10-052-092-10

Query Match
Best Local Similarity 30.38; Score 763; DB 9; Length 547;
Pred No 1 3a-59;
Matches 164; Conservative 66; Mismatches 83; Indels 48; Gaps 7;

QY 127 HNDYICPATNOCTIDKRRKSCQAGRLKCYEWVWVWVCSRRER-----COY 174
DB 149 HNYI-----IADKRNKIKAKFKIRHIFEGSVLADHYGQNTPIGDPVLLPDNY 201
QY 174 KLVKRSKADPQ-----IFAGKAKSGSGHAPPR-----PILLLA 209
DB 202 LSTGSAISKDNEPRIMVILFFVTAABMPSASGAPPAANIWPQGLMIFKPKVCEATG 261
QY 210 LSPDQVLTLEAFPHVLSR--PSAPFTASMMSTLKALKEIVHIMSKAKKIPGV 267
DB 242 LADQVVSALLDAPP--IIVSEYDPTPEFSASMMGLTNLAPELVHIMNAKRVPGV 320
QY 268 FLSFPGVRIIFSGMMFVIMGIMWPSIDHP--KLIFAPPLVILPDGKVEGTEEPML 327

```

```

DB 60 DLTIDVHLLERAMLEFLIMGLVWRSMPKCLIFAPMLIDPNSGKVEGVEIFDML 119
QY 328 LATSPREPLKIQKREYICVKAMILLNSMYVIVAT--GDASSPKLAHLNAVTVLAV 386
DB 120 LATSPREPMNLQEEFVTLKSTILNSVYTFLSSTKSLPKKHIIHVLTKITDTLH 179
QY 387 VIAKSGISSQOOSKRLANLMLLSHYBANSKGMHELLNMKKKNVYVYDILLMLNAH 446
DB 180 LMAKATLIGQGHRLAQIILISHIRMSNKGMEHLVSMKKKNVYVYDILLMLNAH 289
QY 447 L 447
DB 501 L 501

RESULT 23
US-09-853-033-4
Sequence 4; Application US/09853033
Patent No. US20020100068A1
GENERAL INFORMATION:
APPLICANT: CHAMBER, PIERRE
APPLICANT: METZGER, DANIEL
TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
FILE REFERENCE: 065601/9222
CURRENT APPLICATION NUMBER: US/09/853,033
CURRENT FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: FR 00/12570
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent ver. 2.1
SEQ ID NO 4
LENGTH: 660
TYPE: PRI
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence; Chimera sequence
US-09-853-033-4

Query Match
Best Local Similarity 30.38; Score 763; DB 10; Length 660;
Pred No 1 7e-59;
Matches 143; Conservative 57; Mismatches 38; Indels 4; Gaps 4;

QY 259 ALSTFGLVLTLEAFPHVLSR--PSAPFTASMMSTLKALKEIVHIMSKAKKIPGV 266
DB 374 LADQVVSALLDAPP--IIVSEYDPTPEFSASMMGLTNLAPELVHIMNAKRVPGV 432
QY 432 VELSFGVRLIESGMMFVIMGIMWPSIDHP--KLIFAPPLVILPDGKVEGTEEPML 426
DB 433 VDLTFHVVHLLERAMLEFLIMGLVWRSMPKCLIFAPMLIDPNSGKVEGVEIFDML 492
QY 492 LLAISPEPLKIQKREYICVKAMILLNSMYVIVAT--GDASSPKLAHLNAVTVLAV 485
DB 493 LLAISPEPMNLQEEFVTLKSTILNSVYTFLSSTKSLPKKHIIHVLTKITDTLH 552
QY 552 VIAKSGISSQOOSKRLANLMLLSHYBANSKGMHELLNMKKKNVYVYDILLMLNAH 445
DB 553 LMAKATLIGQGHRLAQIILISHIRMSNKGMEHLVSMKKKNVYVYDILLMLNAH 612
QY 446 VL 447
DB 613 KL 614

RESULT 24
US-09-903-876-1
Sequence 1; Application US/09903876
Patent No. US20020072587A1
GENERAL INFORMATION:
APPLICANT: Xu, Zhao-Bao
APPLICANT: Somers, William

```

1 APPLICANT: Akopian, Tatou
 2 APPLICANT: Istiaq, Chai Lai
 3 APPLICANT: Imdad, Raymond
 4 TITLE OF INVENTION: CRYSTAL STRUCTURE OF ESTROGEN RECEPTOR-B COMPLEX
 5 TITLE OF INVENTION: AND USES THEREOF
 6 FILE REFERENCE: 9717479
 7 CURRENT APPLICATION NUMBER: US/09/903,876
 8 PRIOR FILING DATE: 2001-07-11
 9 PRIOR APPLICATION NUMBER: US 60/217,834
 10 PRIOR FILING DATE: 2000-07-12
 11 NUMBER OF SEQ ID NOS: 2
 12 SOFTWARE: Patented version 4.0
 13 SEQ ID NO: 1
 14 LENGTH: 243
 15 TYPE: PRT
 16 ORGANISM: Homo sapiens
 17 US: 09 903 876 1

Query Match 40.2%; Score 762; DB 10; Length 243;
 Best Local Similarity 59.1%; Pred. No. 4,96,60;
 Matches 143; Conservative 57; Mismatches 48; Indels 4; Gaps 4;

QY 209 ALSPEDVITLLEPRPHVLSR--PSAPTEASMMMSLTKLADKELVHMSMAKTVPR 256
 DB 4 SLADQVNSALLDDEP-ILYSEYDTPRPSEASMMGLTNADRELVHIMAKRVPR 61
 QY 267 VELSLPQVRLLESQWMEVIMMIMMSIDHPRKLFAPQVLDHRECKVETSLLEPRM 426
 DB 62 VDLTHQVHLLEFVAMEILMLGVMSMEHPKLFAPNILLDRGCKVEMSVLEPRM 121
 QY 427 LLAISPRRELKQHKRYLVKAMGLNSMYPLVAT QVASSPKLAILNAVITALV 685
 DB 122 LLAISPRKMNIGDEPVVTKSTLNSVYLFSTLSLEKIDHIVLDRITDIL 181
 QY 486 VWIARSLSSQGSQMRANILMLSHVIRASNGMEHLNMRKNVVPYDILLENLNNH 445
 DB 182 HLMKAKGLTQGOHQRLAQLLLSHIRMSNGMEHLNMRKNVVPYDILLENLNNH 241
 QY 446 VL 447
 DB 242 RL 243

RESULT 25
 US: 09 853 033 8
 1 Sequence 8, Application US/09853033
 2 Patent No. US2002010008A1
 3 GENERAL INFORMATION:
 4 APPLICANT: CHAMMON, PIERRE
 5 APPLICANT: METZGER, DANIEL
 6 TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
 7 FILE REFERENCE: 065691/0222
 8 CURRENT APPLICATION NUMBER: US/09/853,033
 9 PRIOR FILING DATE: 2001-05-11
 10 PRIOR APPLICATION NUMBER: FR 00/12970
 11 PRIOR FILING DATE: 2000-10-03
 12 NUMBER OF SEQ ID NOS: 14
 13 SOFTWARE: Patented ver. 2.1
 14 SEQ ID NO: 8
 15 LENGTH: 660
 16 TYPE: PRT
 17 ORGANISM: Artificial sequence
 18 FEATURE:
 19 OTHER INFORMATION: Description of Artificial Sequence: chimeric sequence
 20 OTHER INFORMATION: Homosapiens bacteriophage p1
 US: 09 853 033 8

Query Match 40.1%; Score 760; DB 10; Length 660;
 Best Local Similarity 58.7%; Pred. No. 4,10,59;
 Matches 142; Conservative 57; Mismatches 49; Indels 4; Gaps 4;
 QY 209 ALSPEDVITLLEPRPHVLSR--PSAPTEASMMMSLTKLADKELVHMSMAKTVPR 256
 DB 4 SLADQVNSALLDDEP-ILYSEYDTPRPSEASMMGLTNADRELVHIMAKRVPR 61

DB 474 SLADQVNSALLDDEP-ILYSEYDTPRPSEASMMGLTNADRELVHIMAKRVPR 432
 QY 267 VELSLPQVRLLESQWMEVIMMIMMSIDHPRKLFAPQVLDHRECKVETSLLEPRM 426
 DB 433 VDLTHQVHLLEFVAMEILMLGVMSMEHPKLFAPNILLDRGCKVEMSVLEPRM 492
 QY 427 LLAISPRRELKQHKRYLVKAMGLNSMYPLVAT QVASSPKLAILNAVITALV 685
 DB 493 LLAISPRKMNIGDEPVVTKSTLNSVYLFSTLSLEKIDHIVLDRITDIL 552
 QY 486 VWIARSLSSQGSQMRANILMLSHVIRASNGMEHLNMRKNVVPYDILLENLNNH 445
 DB 553 HLMKAKGLTQGOHQRLAQLLLSHIRMSNGMEHLNMRKNVVPYDILLENLNNH 612
 QY 446 VL 447
 DB 613 RL 614

RESULT 26
 US: 09 853 033 6
 1 Sequence 6, Application US/09853033
 2 Patent No. US2002010008A1
 3 GENERAL INFORMATION:
 4 APPLICANT: CHAMMON, PIERRE
 5 APPLICANT: METZGER, DANIEL
 6 TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
 7 FILE REFERENCE: 065691/0222
 8 CURRENT APPLICATION NUMBER: US/09/853,033
 9 PRIOR FILING DATE: 2001-05-11
 10 PRIOR APPLICATION NUMBER: FR 00/12970
 11 PRIOR FILING DATE: 2000-10-03
 12 NUMBER OF SEQ ID NOS: 14
 13 SOFTWARE: Patented ver. 2.1
 14 SEQ ID NO: 6
 15 LENGTH: 660
 16 TYPE: PRT
 17 ORGANISM: Artificial sequence
 18 FEATURE:
 19 OTHER INFORMATION: Description of Artificial Sequence: chimeric sequence
 20 OTHER INFORMATION: Homosapiens bacteriophage p1
 US: 09 853 033 6

Query Match 29.8%; Score 751; DB 10; Length 660;
 Best Local Similarity 58.3%; Pred. No. 1,96,58;
 Matches 141; Conservative 57; Mismatches 40; Indels 4; Gaps 3;
 QY 209 ALSPEDVITLLEPRPHVLSR--PSAPTEASMMMSLTKLADKELVHMSMAKTVPR 256
 DB 474 SLADQVNSALLDDEP-ILYSEYDTPRPSEASMMGLTNADRELVHIMAKRVPR 432
 QY 267 VELSLPQVRLLESQWMEVIMMIMMSIDHPRKLFAPQVLDHRECKVETSLLEPRM 426
 DB 433 VDLTHQVHLLEFVAMEILMLGVMSMEHPKLFAPNILLDRGCKVEMSVLEPRM 492
 QY 427 LLAISPRRELKQHKRYLVKAMGLNSMYPLVAT QVASSPKLAILNAVITALV 685
 DB 493 LLAISPRKMNIGDEPVVTKSTLNSVYLFSTLSLEKIDHIVLDRITDIL 552
 QY 486 VWIARSLSSQGSQMRANILMLSHVIRASNGMEHLNMRKNVVPYDILLENLNNH 445
 DB 553 HLMKAKGLTQGOHQRLAQLLLSHIRMSNGMEHLNMRKNVVPYDILLENLNNH 612
 QY 446 VL 447
 DB 613 RL 614
 RESULT 27
 US: 10 444 583 1
 1 Sequence 1, Application US/10344583


```

1 APPLICANT: Xiao, Jia Hao
2 APPLICANT: Kusari, Jyotiimoy
3 APPLICANT: Chandaratna, Koshanitha A.
4 TITLE OF INVENTION: Methods of Screening for Compounds That
5 FILE REFERENCE: P-AR 4681
6 TITLE OF INVENTION: Modulate Hormone Receptor Activity
7 CURRENT APPLICATION NUMBER: US/09/922-226
8 PRIOR APPLICATION NUMBER: US 60/294,797
9 PRIOR FILING DATE: 2001-04-18
10 NUMBER OF SEQ ID NOS: 191
11 SOFTWARE: FASTSEQ for Windows Version 4.0
12 SEQ ID NO 2
13 LENGTH: 533
14 TYPE: PRT
15 ORGANISM: Homo sapiens
16 US-09-922-226-2

```

```

Query Match 18.2% Score 458.5; Dh 9; Length 533;
Best Local Similarity 40.2%; Pred. No. 1,96-32;
Matches 140; Conservative 63; Mismatches 163; Indels 97; Gaps 17;

```

```

01 9 VTNI FGGGPGGRTSPNVLWPTPGHLSPLVVRHLSHLYARPOKSPWCEARSLHLLPVN- 67
02 146 ISSMGSGGRLPPVAP-----PG-----FSQVSSVQINSTV---SLPQGG 172
03 68 -PPTKPKVSSNR--CASVTPGSKEDAHFVAVCSVASCYVAVCSVAVVAVV 121
04 173 SPPPIVPRVIVVWGLHPPPOGPNZK--RICAICGPPSSGKHVAVVSPVGVGFF 230
05 122 PGLGCHNIVYCPATNLTTHKPKKSGVAVRPKGVAVVAVVAVVAVVAVVAVV 181
06 241 FTRKRLTYSPVWPKSTVPRP--GVVGVGVGVGVGVGVGVGVGVGVGVGVGV 282
07 182 AHGVPVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 229
08 283 KKKK--GGGPGG--APPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 322
09 240 -SPPVATTESSMMNSTKLAKELVIMISMAKRTGVFVSLFVGLFVW 282
10 423 GGTGSSSSSPNIVT-----NICQAKMKLFLVEMARFTRSSLPVAVV111PAGW 376
11 423 GGTGSSSSSPNIVT-----NICQAKMKLFLVEMARFTRSSLPVAVV111PAGW 376
12 283 MEVIMMGLMWSSTHHPKRLIPADLVLDROGCVGDTTLPIMLLA-TTSRELELKH 341
13 477 NEETLASHSHSTIVVWGLHAGVAVHVPNSASHA-GVCAFLPVPVTEI VSKRPMEMK 435
14 442 KPVGVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 401
15 436 TELGTVLAVLPSIGKLTLEHFFPKLIGVPIDTLEMLLA 528
16 402 LANLLMLLSHVHANSKGMELLMKCKNVVPVVDLEMLLA 444
17 486 FAKLLLPALPSIGKLTLEHFFPKLIGVPIDTLEMLLA 528

```

```

RESULT 33
US-10-098-184-2
1 Sequence 2; Application US/10098184
2 Publication No. US2003010533A1
3 GENERAL INFORMATION:
4 APPLICANT: Plahl, Mannus
5 APPLICANT: Tachdjian, Catherine
6 APPLICANT: Al-Shamma, Hussein A
7 APPLICANT: Fanjul, Andrea
8 APPLICANT: Pleyner, David P. M.
9 APPLICANT: Spruce, Kyle W.
10 APPLICANT: Fine, Richard
11 APPLICANT: Zapf, James W.
12 TITLE OF INVENTION: RXR ACTIVATING MOLECULES
13 FILE REFERENCE: 14099, 001602
14 CURRENT APPLICATION NUMBER: US/10/098,184
15 CURRENT FILING DATE: 2002-03-08
16 PRIOR APPLICATION NUMBER: 60/274,342

```

```

1 PRIOR FILING DATE: 2001-03-08
2 NUMBER OF SEQ ID NOS: 3
3 SOFTWARE: FASTSEQ for Windows Version 4.0
4 SEQ ID NO 2
5 LENGTH: 533
6 TYPE: PRT
7 ORGANISM: Artificial Sequence
8 FEATURE INFORMATION: Description of Artificial Sequence: No. US2003010533A1
9 OTHER INFORMATION: synthetic construct
10 US-10-098-184-2

```

```

Query Match 18.2% Score 458.5; Dh 9; Length 533;
Best Local Similarity 40.2%; Pred. No. 1,96-32;
Matches 140; Conservative 63; Mismatches 163; Indels 97; Gaps 17;

```

```

01 9 VTNI FGGGPGGRTSPNVLWPTPGHLSPLVVRHLSHLYARPOKSPWCEARSLHLLPVN- 67
02 146 ISSMGSGGRLPPVAP-----PG-----FSQVSSVQINSTV---SLPQGG 172
03 68 -PPTKPKVSSNR--CASVTPGSKEDAHFVAVCSVASCYVAVCSVAVVAVVAVV 121
04 173 SPPPIVPRVIVVWGLHPPPOGPNZK--RICAICGPPSSGKHVAVVSPVGVGFF 230
05 122 PGLGCHNIVYCPATNLTTHKPKKSGVAVRPKGVAVVAVVAVVAVVAVVAVV 181
06 241 FTRKRLTYSPVWPKSTVPRP--GVVGVGVGVGVGVGVGVGVGVGVGVGVGV 282
07 182 AHGVPVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 229
08 283 KKKK--GGGPGG--APPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 322
09 240 -SPPVATTESSMMNSTKLAKELVIMISMAKRTGVFVSLFVGLFVW 282
10 423 GGTGSSSSSPNIVT-----NICQAKMKLFLVEMARFTRSSLPVAVV111PAGW 376
11 423 GGTGSSSSSPNIVT-----NICQAKMKLFLVEMARFTRSSLPVAVV111PAGW 376
12 283 MEVIMMGLMWSSTHHPKRLIPADLVLDROGCVGDTTLPIMLLA-TTSRELELKH 341
13 477 NEETLASHSHSTIVVWGLHAGVAVHVPNSASHA-GVCAFLPVPVTEI VSKRPMEMK 435
14 442 KPVGVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 401
15 436 TELGTVLAVLPSIGKLTLEHFFPKLIGVPIDTLEMLLA 528
16 402 LANLLMLLSHVHANSKGMELLMKCKNVVPVVDLEMLLA 444
17 486 FAKLLLPALPSIGKLTLEHFFPKLIGVPIDTLEMLLA 528

```

```

RESULT 34
US-09-922-226-3
1 Sequence 3; Application US/09922226
2 Publication No. US2003007766A1
3 GENERAL INFORMATION:
4 APPLICANT: Zhao, Yi
5 APPLICANT: Thacher, Scott M.
6 APPLICANT: Xiao, Jia Hao
7 APPLICANT: Kusari, Jyotiimoy
8 APPLICANT: Chandaratna, Koshanitha A.
9 TITLE OF INVENTION: Methods of Screening for Compounds That
10 FILE REFERENCE: P-AR 4681
11 CURRENT APPLICATION NUMBER: US/09/922-226
12 CURRENT FILING DATE: 2002-01-09
13 PRIOR APPLICATION NUMBER: US 60/284,797
14 PRIOR FILING DATE: 2001-04-18
15 NUMBER OF SEQ ID NOS: 191
16 SOFTWARE: FASTSEQ for Windows Version 4.0
17 SEQ ID NO 3
18 LENGTH: 463
19 TYPE: PRT
20 ORGANISM: Homo sapiens
21 US-09-922-226-3

```

```

Query Match      17.4M: Score 440; DB 9; Length 464;
Best Local Similarity 42.1M: Prod. No. 6, 90-41;
Matches 116; Conservative 56; Mismatches 157; Indels 42; Gaps 8;

DB 146 KORNRCUYCRKQKTLWKKRFAVGEF-----GKSRERASAFABA-----TSGHEMD 245
DB 204 RELLDLMSFPOVLTLLFAEPBVLTSPASAFTHASMMSTTKLANKEIWMISWAK 262
DB 246 VERTEFA---ELAVEPKESTVGMNMNSTNGVT-----NICHADKQGLFTLYWAKR 286
DB 264 ITCGVTSITGVVLTIVWVWVIMMIMMPSITHEKTLFVAVLVIGERECVYVLTLE 422
DB 287 ITHSDLTLEOVLTLMKNNELLASPSRSVSDVGLTLATGLIHHSASHA GYVS 345
DB 423 IEMMLA-TSRPEELKIOREYLCYKAMILLNSWVPTVATODASSKTLAHJNAV 381
DB 446 ITHVLTLEVSKMOMKSHLOGLALVLENDKAKLSNSEVELLEKRYATLEAT 405
DB 482 IATWVLTAKSNTSSGQSMPLANIMLSIVRIASNKMEHLNMRKRVVIVYVLTLE 441
DB 496 KQ-----KYPGQGRFANILRLALAKSLGKLEHLEFFKLTGDTLIDPLEM 455
DB 442 L 442
DB 456 L 456

RESULT 35
DB 10 098 184 3
Sequence 3; Application US/10098184
Publication No. US2001010544A1
GENERAL INFORMATION:
APPLICANT: Pradit, Magnus
APPLICANT: Thachiatu, Catherine
APPLICANT: Al Shammou, Hussein A.
APPLICANT: Fanchou, Andrew
APPLICANT: Pleginet, David P. M.
APPLICANT: Spruce, Lytle W.
APPLICANT: Fries, Richard
APPLICANT: Zapf, James W.
TITLE OF INVENTION: PCR ACTIVATING MODULATORS
FILE REFERENCE: 1099,001602
CURRENT APPLICATION NUMBER: US/10098, 184
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: 60/274, 342
PRIOR FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 464
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence; No. US2001010544A1;
US 10-098 184 3

Query Match      17.4M: Score 440; DB 9; Length 464;
Best Local Similarity 42.1M: Prod. No. 6, 90-41;
Matches 116; Conservative 56; Mismatches 157; Indels 42; Gaps 8;

DB 146 KORNRCUYCRKQKTLWKKRFAVGEF-----GKSRERASAFABA-----TSGHEMD 245
DB 204 RELLDLMSFPOVLTLLFAEPBVLTSPASAFTHASMMSTTKLANKEIWMISWAK 262
DB 246 VERTEFA---ELAVEPKESTVGMNMNSTNGVT-----NICHADKQGLFTLYWAKR 286
DB 264 ITCGVTSITGVVLTIVWVWVIMMIMMPSITHEKTLFVAVLVIGERECVYVLTLE 422
DB 287 ITHSDLTLEOVLTLMKNNELLASPSRSVSDVGLTLATGLIHHSASHA GYVS 345
DB 423 IEMMLA-TSRPEELKIOREYLCYKAMILLNSWVPTVATODASSKTLAHJNAV 381
DB 446 ITHVLTLEVSKMOMKSHLOGLALVLENDKAKLSNSEVELLEKRYATLEAT 405
DB 482 IATWVLTAKSNTSSGQSMPLANIMLSIVRIASNKMEHLNMRKRVVIVYVLTLE 441
DB 496 KQ-----KYPGQGRFANILRLALAKSLGKLEHLEFFKLTGDTLIDPLEM 455
DB 442 L 442
DB 456 L 456

RESULT 35
DB 10 098 184 3
Sequence 3; Application US/10098184
Publication No. US2001010544A1
GENERAL INFORMATION:
APPLICANT: Pradit, Magnus
APPLICANT: Thachiatu, Catherine
APPLICANT: Al Shammou, Hussein A.
APPLICANT: Fanchou, Andrew
APPLICANT: Pleginet, David P. M.
APPLICANT: Spruce, Lytle W.
APPLICANT: Fries, Richard
APPLICANT: Zapf, James W.
TITLE OF INVENTION: PCR ACTIVATING MODULATORS
FILE REFERENCE: 1099,001602
CURRENT APPLICATION NUMBER: US/10098, 184
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: 60/274, 342
PRIOR FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 464
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence; No. US2001010544A1;
US 10-098 184 3

```

```

DB 146 KORNRCUYCRKQKTLWKKRFAVGEF-----GKSRERASAFABA-----TSGHEMD 245
DB 204 RELLDLMSFPOVLTLLFAEPBVLTSPASAFTHASMMSTTKLANKEIWMISWAK 262
DB 246 VERTEFA---ELAVEPKESTVGMNMNSTNGVT-----NICHADKQGLFTLYWAKR 286
DB 264 ITCGVTSITGVVLTIVWVWVIMMIMMPSITHEKTLFVAVLVIGERECVYVLTLE 422
DB 287 ITHSDLTLEOVLTLMKNNELLASPSRSVSDVGLTLATGLIHHSASHA GYVS 345
DB 423 IEMMLA-TSRPEELKIOREYLCYKAMILLNSWVPTVATODASSKTLAHJNAV 381
DB 446 ITHVLTLEVSKMOMKSHLOGLALVLENDKAKLSNSEVELLEKRYATLEAT 405
DB 482 IATWVLTAKSNTSSGQSMPLANIMLSIVRIASNKMEHLNMRKRVVIVYVLTLE 441
DB 496 KQ-----KYPGQGRFANILRLALAKSLGKLEHLEFFKLTGDTLIDPLEM 455
DB 442 L 442
DB 456 L 456

RESULT 36
DB 09 922 226 1
Sequence 1; Application US/09922226
Publication No. US2001007766A1
GENERAL INFORMATION:
APPLICANT: Zhao, Yi
APPLICANT: Thachert, Scott M.
APPLICANT: Xiao, Jia Hao
APPLICANT: Kosarli, Jyotiimoy
APPLICANT: Chandraratna, Roshanika A.
TITLE OF INVENTION: Methods of Screening for Compounds That
Modulate Hormone Receptor Activity
FILE REFERENCE: P AR 4681
CURRENT APPLICATION NUMBER: US/09/922, 226
CURRENT FILING DATE: 2002-01-09
PRIOR APPLICATION NUMBER: US 60/284, 797
PRIOR FILING DATE: 2001-04-18
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 462
TYPE: PRT
ORGANISM: Homo sapiens
US 09 922 226 1

Query Match      17.2M: Score 434.5; DB 9; Length 462;
Best Local Similarity 28.9M: Prod. No. 2, 10-40;
Matches 139; Conservative 67; Mismatches 206; Indels 69; Gaps 14;

DB 1 MNSVTPSNVTNIEGSGKQTSNIVLPT--PGLSPVAVHGLSDVA--PQKSPWCE 56
DB 9 ITHSDLTLEOVLTLMKNNELLASPSRSVSDVGLTLATGLIHHSASHA GYVS 345
DB 57 ARSL--EHTLV-NRETLKRVASQNGASVPTVTHSKRP----- 92
DB 68 ITHSDLTLEOVLTLMKNNELLASPSRSVSDVGLTLATGLIHHSASHA GYVS 345
DB 94 ---AHFAGVSYVASYVGVWSTGCAKFAFKSTGGINHYGTATNGTIDKRRKSC 148
DB 128 ASPTKHVATVGDSSSKRHVGVVSCRCCKPFRKIVYTCVQINRQVTLKQQRNG 187
DB 149 VATHKRGVGVWGVVGVSGRKGQYKIVYVGRKSAFQGLDANAKKSTGHAIVRELLD 208
DB 188 QYTKVQKTLAMKRRKAVQGE-----KQKQKQKNGVSTSSANIMV-VERTEFA 248
DB 209 ALSTPQVLTLEA--EYHVLNRSAPVTHASMMSTTKLANKEIWMISWAK 264
DB 249 ELAVEPKESTVGMNMNSTNGVT-----NICHADKQGLFTLYWAKR 287

```

265 GCVLSTLFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 424
DB HESTLFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 424
425 LMLA-TSPRFLK-GRKFTCVKAVL-NESMYELVATG-ANGSSTLARI-NAYTA- 425
DB LMLA-TSPRFLK-GRKFTCVKAVL-NESMYELVATG-ANGSSTLARI-NAYTA- 425
447 DPAVLTIVSPKPMGMVLTGCTPATVLTINRSPKLSHRAVFAVAPVYVANGTATV 447
DB LMLA-TSPRFLK-GRKFTCVKAVL-NESMYELVATG-ANGSSTLARI-NAYTA- 447
484 LVMVIAKSGISGOSGSMRLANLMLSHVHASNKRMPHINMKRKNVAVYVGLLEMLN 444
DB LVMVIAKSGISGOSGSMRLANLMLSHVHASNKRMPHINMKRKNVAVYVGLLEMLN 444
407 -----KYPGQPKFAKLLRLPALRSICLKLHLPEPKLTIDPILTEMEMLE 456
DB 407 -----KYPGQPKFAKLLRLPALRSICLKLHLPEPKLTIDPILTEMEMLE 456
444 A 444
DB 457 A 457

RESULT 47
US-10-098-184-1
Sequence 1, Application US/10098184
Publication No. US20030105333A1
GENERAL INFORMATION:
APPLICANT: Pihl, Magnus
APPLICANT: Tachjian, Catherine
APPLICANT: Al-Shamma, Hussein A.
APPLICANT: Fajulu, Andrea
APPLICANT: Spruce, David P.M.
APPLICANT: Spruce, Lyle W.
APPLICANT: Zapt, James W.
FILE OF INVENTION: RKR ACTIVATING MOLECULES
FILE REFERENCE: 13099, 001602
CURRENT APPLICATION NUMBER: US/10-098-184
PRIOR FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: 60/274,342
PRIOR FILING DATE: 2001-03-08
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 462
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence; No. US20030105333A1-
US-10-098-184-1

Query Match 17.2% Score 434.5 DB 9 Length 462
Best local Similarity 28.9% Prod. No. 2,1e-30
Matches 139 Conservative 67 Mismatches 205 Indels 69 Gaps 14

1 MNYSTFNVNTEGCGPQRTSPNVMT--PCHSLVYHQLSHLYA--PQKSPMP 56
DB 9 LDFSTQVN--SLSTPFGSGMAVSLHPSLGLGSGNGLHSPILSTLSL--INMGFEAN 67
57 AKST--PHTTPV--NPFTFKKVSNGPCASPVTPGSKPD----- 92
DB 57 AKST--PHTTPV--NPFTFKKVSNGPCASPVTPGSKPD----- 92
68 ISSMGHSMGVPTPTPLFSTSPGSSPMNVVSSFTGKPKLITNVIKVAHDSNM 127
DB 68 ISSMGHSMGVPTPTPLFSTSPGSSPMNVVSSFTGKPKLITNVIKVAHDSNM 127
94 -----AHTAVGSGVASYVYVWSTFGKAFKFSVCHNOVGVAINNGTTHNNGPSS 145
DB 94 -----AHTAVGSGVASYVYVWSTFGKAFKFSVCHNOVGVAINNGTTHNNGPSS 145
128 ASFTKHAVALGSGSSGTHGVVVSFGVCFEPFTVKKLITVTF--NMGSTLDFGSGNGL 167
DB 128 ASFTKHAVALGSGSSGTHGVVVSFGVCFEPFTVKKLITVTF--NMGSTLDFGSGNGL 167
145 GATFLAKVYVWAKGSSPFPQVYPLVVPQASVGLHLCATFAFSSGSHAPVRELLIT 208
DB 145 GATFLAKVYVWAKGSSPFPQVYPLVVPQASVGLHLCATFAFSSGSHAPVRELLIT 208
188 GYFVYVGTALMGKMFVAVGFE-----PQKSTGPNVVSIS--SANGEM--VEKLLIA 428
DB 188 GYFVYVGTALMGKMFVAVGFE-----PQKSTGPNVVSIS--SANGEM--VEKLLIA 428
239 AHSPPGVLTILFA-----PFFHVLSPSSAFETFASMMMSLTETACFEVVMKSWAFTE 264
DB 239 AHSPPGVLTILFA-----PFFHVLSPSSAFETFASMMMSLTETACFEVVMKSWAFTE 264
249 FIAFPPKTFVVFANMGIND-----SSPNQVLT-----NIGQAAVKGQFTVFAVAPV 297
DB 249 FIAFPPKTFVVFANMGIND-----SSPNQVLT-----NIGQAAVKGQFTVFAVAPV 297
265 GCVLSTLFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 424
DB 265 GCVLSTLFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 424

268 HPSLSTFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 446
DB HPSLSTFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 446
425 LMLA-TSPRFLK-GRKFTCVKAVL-NESMYELVATG-ANGSSTLARI-NAYTA- 484
DB LMLA-TSPRFLK-GRKFTCVKAVL-NESMYELVATG-ANGSSTLARI-NAYTA- 484
447 DPAVLTIVSPKPMGMVLTGCTPATVLTINRSPKLSHRAVFAVAPVYVANGTATV 447
DB LMLA-TSPRFLK-GRKFTCVKAVL-NESMYELVATG-ANGSSTLARI-NAYTA- 447
484 LVMVIAKSGISGOSGSMRLANLMLSHVHASNKRMPHINMKRKNVAVYVGLLEMLN 444
DB LVMVIAKSGISGOSGSMRLANLMLSHVHASNKRMPHINMKRKNVAVYVGLLEMLN 444
407 -----KYPGQPKFAKLLRLPALRSICLKLHLPEPKLTIDPILTEMEMLE 456
DB 407 -----KYPGQPKFAKLLRLPALRSICLKLHLPEPKLTIDPILTEMEMLE 456
444 A 444
DB 457 A 457

RESULT 48
US-09-965-703-63
Sequence 63, Application US/09965703
Patent No. US20020119521A1
GENERAL INFORMATION:
APPLICANT: Rohm and Haas Company
APPLICANT: Paili, Subba Reddy
APPLICANT: Kapitskaya, Marianna Zinovjevna
APPLICANT: Gress, Dean Ervin
FILE REFERENCE: A010208
CURRENT APPLICATION NUMBER: US/09-965-703
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/191,355
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/269,799
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/205,050
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 63
LENGTH: 467
TYPE: PRT
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: No. US20020119521A1 sequence
US-09-965-703-63

Query Match 17.2% Score 434. DB 10 Length 467
Best local Similarity 28.9% Prod. No. 2,4e-30
Matches 140 Conservative 67 Mismatches 206 Indels 72 Gaps 14

1 MNYSTFNVNTEGCGPQRTSPNVMT--PCHSLVYHQLSHLYA--PQKSPMP 68
DB 9 LDFSTQVN--SLSTPFGSGMAVSLHPSLGLGSGNGLHSPILSTLSL--INMGFEAN 67
53 FVFAKSL--PHTTPV--NPFTFKKVSNGPCASPVTPGSKPD----- 92
DB 53 FVFAKSL--PHTTPV--NPFTFKKVSNGPCASPVTPGSKPD----- 92
68 ISSMGHSMGVPTPTPLFSTSPGSSPMNVVSSFTGKPKLITNVIKVAHDSNM 128
DB 68 ISSMGHSMGVPTPTPLFSTSPGSSPMNVVSSFTGKPKLITNVIKVAHDSNM 128
94 -----AHTAVGSGVASYVYVWSTFGKAFKFSVCHNOVGVAINNGTTHNNGPSS 144
DB 94 -----AHTAVGSGVASYVYVWSTFGKAFKFSVCHNOVGVAINNGTTHNNGPSS 144
128 ASFTKHAVALGSGSSGTHGVVVSFGVCFEPFTVKKLITVTF--NMGSTLDFGSGNGL 167
DB 128 ASFTKHAVALGSGSSGTHGVVVSFGVCFEPFTVKKLITVTF--NMGSTLDFGSGNGL 167
145 GATFLAKVYVWAKGSSPFPQVYPLVVPQASVGLHLCATFAFSSGSHAPVRELLIT 204
DB 145 GATFLAKVYVWAKGSSPFPQVYPLVVPQASVGLHLCATFAFSSGSHAPVRELLIT 204
188 GYFVYVGTALMGKMFVAVGFE-----PQKSTGPNVVSIS--SANGEM--VEKLLIA 428
DB 188 GYFVYVGTALMGKMFVAVGFE-----PQKSTGPNVVSIS--SANGEM--VEKLLIA 428
239 AHSPPGVLTILFA-----PFFHVLSPSSAFETFASMMMSLTETACFEVVMKSWAFTE 260
DB 239 AHSPPGVLTILFA-----PFFHVLSPSSAFETFASMMMSLTETACFEVVMKSWAFTE 260
249 FIAFPPKTFVVFANMGIND-----SSPNQVLT-----NIGQAAVKGQFTVFAVAPV 288
DB 249 FIAFPPKTFVVFANMGIND-----SSPNQVLT-----NIGQAAVKGQFTVFAVAPV 288
265 GCVLSTLFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 424
DB 265 GCVLSTLFDVRLTFCGMMVLMGMLMMSLDRPKLTAPDVLTFFDQKVEVLELE 424

Search completed: June 16, 2003, 13:41:51
Job time : 53.9571 secs

•
•
•

RatMyoma, P. Finch, H. Sasaki, 1993; H. Kawan, H. Shimizu, H. Hoshimoto, S. Kato
 Hirobumi, H. Hoshino, Commun. 245, 142-147, 1998
 A:Title: A novel isoform of rat estrogen receptor beta with 18 amino acid insertion in C
 A:Reference number: JMW0046; MUID:98262942; PMID:9600083
 A:Accession: JMW0046
 A:Molecule type: mRNA
 A:Residues: 1504 - MAR
 A:Cross-reference: DDBJ:AA012721
 Comment: This protein functions as a negative regulator of estrogen action.
 Comment: This protein receptor with transforming protein homology
 F102-481/yeast-yeast transforming protein homology -ERBA

Query Match 48.9% Score 2343; Dh 2; Length 589
 best Local Similarity 85.7% Prod. No. 1c 179;
 Matches 421; Conservative 23; Mismatches 30; Indels 18; Gaps 1;

```

UY 1 MNYSPSNVTNPRPSPPTSPNVPPTPHNHSPIVWHPQSHLYAFPKSKVFAST 40
    ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9 MNYSPVSTSNLHGVPGSTSPNVLMPISGLSPHATHGSLYAFPKSPWFAKSL 68

UY 61 EHLVYNNRETLKRVSNCFASVTPDPSKPAHFVAVSGVSYHYGVWSTETKAF 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69 EHLVYNNRETLKRLKSLSSASVSPNAKPAHFVAVSGVSYHYGVWSTETKAF 128

UY 121 KRSLGHNHYTCATNACFLPNNPKSGACAFKQCEYGVWGVGSGSPRPGCYLVPQR 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 129 KRSLGHNHYTCATNACFLPNNPKSGACAFKQCEYGVWGVGSGSPRPGCYLVPQR 188

UY 181 SAEQHLGACAKKPSGSHAPVPEELLDAISPEQLVTLLEAPPHVLSRSPADFEAS 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 189 SSSEGVHGLSKARKNGHAPVKEELLSTLSPEQLVTLLEAPPHVLSRSPADFEAS 248

UY 241 MMSSTTKIAKRELVHMSNAKKIPGFVPSIFGVYPLLESCMEVLMIMPSIDPQR 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 MMSSTTKIAKRELVHMSNAKKIPGFVPSIFGVYPLLESCMEVLMIMPSIDPQR 308

UY 301 LEPAPDIVD-----PFEKCVGSLLETFMLLATISPREELQHK 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 309 LEPAPDIVDSESDPHWVAQMKSAAPRPGKGVGILETFMLLATISPREELQHK 368

UY 343 EYLCVAMILLNSMPVATADASSKLAHTNAVTDAVWYIANSGLSSQVSMHL 402
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 369 EYLCVAMILLNSMPVATADASSKLAHTNAVTDAVWYIANSGLSSQVSMHL 428

UY 403 ANLLMLSHVRIASNGMELLMKKNVVPYDILLMLNAHVHKGKSSISGSPSPA 462
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 429 ANLLMLSHVRIASNGMELLMKKNVVPYDILLMLNAHVHKGKSSISGSPSPA 488

UY 463 EDSKSKRSONPQSQ 477
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 EDSKSKRSONPQSQ 503
```

RESULT 4

ORIGIN

estrogen receptor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 06-Mar-1992 #sequence_revision 14-Jul-1994 #text_change 22-Jun-1999
 C:Accession: A40914; MUID:88418621; PMID:2901032

R:Maxwell, B.L.; McConnell, D.P.; Connolly, G.M.; Schultz, T.Z.; Greene, G.L.; O'Malley,
 Mol. Endocrinol. 1, 25-35, 1987

A:Title: Structural organization and regulation of the chicken estrogen receptor.
 A:Reference number: A40914; MUID:88418621; PMID:2901032

A:Accession: A40914
 A:Status: not compared with conceptual translation

A:Molecule type: mRNA
 A:Residues: 1-589 -MAX-

K:Krusty, A.; Green, S.; Argos, P.; Kumar, V.; Walter, P.; Bornert, J.M.; Chambon, P.
 EMBO J. 5, 891-897, 1986

A:Title: The chicken estrogen receptor sequence: homology with v-erbA and the human oes
 A:Reference number: S07192; MUID:86247578; PMID:3755102

A:Accession: S07192
 A:Status: preliminary

A:Molecule type: mRNA
 A:Residues: 1-256; F, 258-589 -KRV>
 A:Cross-reference: EMBL:X04805; NID:q64378; PDB:1CAA2743; 1; PID:q63480
 Comment: The steroid hormones and their receptors are involved in the regulation of
 Comment: In the absence of ligand, steroid hormone receptors are thought to be weak
 complex appears to recognize discrete DNA sequences upstream of transcriptional start
 C:Superfamily: estrogen receptor, with transforming protein homology
 C:Keywords: DNA binding, nucleus, phosphoprotein, steroid binding, steroid hormone re
 F1-174/Domain: amino terminal; NH2
 F124 265/Domain: DNA binding #status predicted <DNA>
 F127 450/Domain: cDNA transforming protein homology -ERBA
 F177-200/Region: zinc finger CCCC motif
 F232 295/Region: zinc finger CCCC motif
 F250 265/Region: nuclear location signal
 F294 946/Domain: steroid binding #status predicted <STR>
 F119,182,196,199/Binding site: zinc (Cys) #status predicted
 F215,231,234/Binding site: zinc (Cys) #status predicted
 F120,209/Binding site: phosphate (Ser) (revariant) #status predicted

Query Match 48.6% Score 1226; Dh 1; Length 589;
 best Local Similarity 50.9% Prod. No. 1c 89;
 Matches 251; Conservative 84; Mismatches 112; Indels 46; Gaps 11;

```

UY 22 SPNVIMPTGSLHPLVYVIRQLSHYAFPKSPWFAKSLHPLV-----NRETL 71
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 100 SPNVVGTATQGISPTIHHSQVVPYIPRPGSP-----IMPAAPATYPPSSINRHSI 155

UY 72 KRVVSNPAPVTPDPSKPAHFVAVSGVSYHYGVWSTETKAFKSLGHNHY 131
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 156 KRVVSNPAPVTPDPSKPAHFVAVSGVSYHYGVWSTETKAFKSLGHNHY 214

UY 132 CPATNCTTGNPKSGACAFKQCEYGVWGVGSGSPRPGCYLVPQRSAHOLHCKG 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 CPATNCTTGNPKSGACAFKQCEYGVWGVGSGSPRPGCYLVPQRSAHOLHCKG 274

UY 192 AKPSDCHAFV-----ALSPGCVTLLEAPPHVLSRSPADFEA 239
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 275 AKPSDCHAFV-----ALSPGCVTLLEAPPHVLSRSPADFEA 334

UY 240 SSMMSSTTKIAKRELVHMSNAKKIPGFVPSIFGVYPLLESCMEVLMIMPSIDPQR 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 335 SSMMSSTTKIAKRELVHMSNAKKIPGFVPSIFGVYPLLESCMEVLMIMPSIDPQR 394

UY 300 KLEFAPDIVDPSGCVGSLLETFMLLATISPREELQHKRETVAKAMILLNSMP 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 395 KLEFAPDIVDPSGCVGSLLETFMLLATISPREELQHKRETVAKAMILLNSMP 454

UY 360 LVTAIDADSSKLAHTNAVTDAVWYIANSGLSSQVSMHLMLSHVRIASNG 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 455 LVTAIDADSSKLAHTNAVTDAVWYIANSGLSSQVSMHLMLSHVRIASNG 514

UY 419 GMEHLLMKKKNVVPYDILLMLNAHVHKGKSSISGSPSPA-----GSAE----- 463
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 515 GMEHLLMKKKNVVPYDILLMLNAHVHKGKSSISGSPSPA-----GSAE----- 574

UY 464 ---DSKSKRSON 473
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 575 ---DSKSKRSON 587
```

RESULT 5

ORIGIN

estrogen receptor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 28-Feb-1992 #sequence_revision 14-Jul-1994 #text_change 15-Jun-1999
 C:Accession: A40907

R:Weller, L.S.; Lew, D.; Shapiro, D.J.
 Mol. Endocrinol. 1, 355-362, 1987

A:Title: The Xenopus laevis estrogen receptor: sequence homology with human and avian
 A:Reference number: A40907; MUID:50311927; PMID:3274094

A:Accession: A40907
 A:Molecule type: mRNA
 A:Residues: 1-586 -WET>

Species: Mus musculus (house mouse)
 Date: 14-Mar-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
 Accession: S58087
 R:Pettersson, K.; Svensson, K.; Mattsson, P.; Carlsson, B.; Ohlsson, B.; Bergstrom, A.
 Submitted to the EMBL Data Library, July 1995
 A:Description: Expression of a novel member of estrogen response element-binding protein
 A:Reference number: S58087
 A:Accession: S58087
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-433 <PEP>
 A:Cross-references: EMBL: X89594, NID: q14855, PDB: 1CA6755.1, PDB: q14856
 C:Keywords: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: steroid hormone receptor; zinc finger
 F:101-352/Domain; erba transforming protein homology <ERBA>

Query Match 24.58; Score 617; DB 2; Length 433;
 Best local similarity 38.79; Pred. No. 2, 30-41;
 Matches 139; Conservative 76; Mismatches 108; Indels 36; Gaps 10.

QY 96 CAACSDYASGTHGVASCCEGCAKAFKRSIQGHNIYICATNCTIDKAPKSCAGRLRK 155
 DB 103 CACVCHASGVHGVASCPAKAFKRTIDGNIYVCPATINEFTIKPKKSCAGCPKPV 162
 QY 156 CTEVGVWVKQNSP--PPGCVPLVPPSPADE---GLHCAKAKKSGCHADVPPELILA 209
 DB 163 CLKVMIKRVVPIPVPPSPGCKYKRRHLSFNSTYINLPSIFAKET-----LTSVSNL 216
 QY 210 LSHPEGLVLTLLAPPHVILSRSDAPETASMMSTIKLAKELVHMSMAKKIPGEVL 269
 DB 217 LSEVQ---DKLVAMPN-----DPEEDIKALTLCEIADPELVLMKHHIPGSL 267
 QY 270 STEFVPLIFSCWNEVLMKIMKPSIDHPKELIFNDVILPDEKCPCELLIFLIMLA 329
 DB 268 TLIDQMSITQSNAMPTITCIYKPSIPIYKATAYEDYIMREHSPVYCLDQIPALIQ 324
 QY 330 TTSRFEELKIQKREYLYKAMILLNS-SMYPELVATIDQSSKRIALLIN-AVPIALVW 387
 DB 327 LVPYFKRIKVEKEFEFTKALAINSTMY-----LENLVAQKICVDLIFKALSHVEL- 379
 QY 384 IAKSSTISSGQSMPIANIMLSHVPHASNKCHELLNMCANVDPVDTLLLENNALV 446
 DB 380 -----SQRDEPRKAKRIWTLPIIQTAKAVGHFYSVKIQKVMKRI FLEML EAVV 433

RESULT 16
 A29345
 Steroid hormone receptor ERK1 precursor - human
 N:Alternative names: estrogen-related receptor
 C:Species: Homo sapiens (man)
 C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 20-Sep-1999
 A:Accession: A29345; A19074
 R:Giguere, V.; Yand, N.; Segall, P.; Evans, P. M.
 Nature 331, 91-94, 1988
 A:Title: Identification of a new class of steroid hormone receptors.
 A:Reference number: A29345; MIM: 6812545; FMT: 3257297
 A:Accession: A29345
 A:Molecule type: mRNA
 A:Residues: 1-521 <GIG>
 A:Cross-references: EMBL: X51416, NID: q46608, PDB: 1CA6755.1, PDB: q46609; EMBL: Y00290
 K:Miller, S.R.; Kraus, P. T.; Zuo, F.; Murray, F. E.; Lott, P. J.; Murray, T. F.
 Genes Dev. 7, 2296-2319, 1993
 A:Title: SV40 early-late switch involves titration of cellular transcription factors
 A:Reference number: A49074; MIM: 94040741; PMID: 8224847
 A:Accession: A49074
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 166-169, X, 171-173 <WIL>
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation; z
 F:174-434/Domain; erba transforming protein homology <ERBA>
 F:176-196/Region; zinc finger
 F:212-216/Region; zinc finger

Query Match 23.44; Score 594.5; DB 2; Length 521;
 Best local similarity 49.68; Pred. No. 1, 80-89;
 Matches 149; Conservative 54; Mismatches 142; Indels 41; Gaps 9

QY 86 GPG-----SKRIAMHCAVSGYASGVHGVASCCEGCAKAFKRSIQGHNIYICATN 146
 DB 157 GPGGQGGKLVISIRPKCLVGVASGVHGVASCPAKAFKRTIDGNIYVCPATINEFTIK 216
 QY 137 CTEVGVWVKQNSP--PPGCVPLVPPSPADE---GLHCAKAKKSGCHADVPPELILA 190
 DB 217 CTEVGVWVKQNSP--PPGCVPLVPPSPADE---GLHCAKAKKSGCHADVPPELILA 276
 QY 191 KARSCHAPREVELLIDLSPEQVLTLEMPHHV-LISRSAPITEASMMSTIKLA 249
 DB 277 PLAVAG--PKTAAVNA-----LVSHLLVPEKIVAMPDPAQDCHLPAVATIDLP 329
 QY 250 FRELVMISMAKKIPGEVLSFDQVRLKSCWMEVLMKIMKPSIDHPKELIFNDVILP 309
 DB 330 DEIVVITSMKSTIPGSSLSIDQMSVIGSWMEVILVAGVAPSLIPDIEAFARDELVL 389
 QY 410 DHPGKVPESLIFEDMLATTSPPEELKIQKREYLYKAMILLNS-SMYPELVATIDQADS 369
 DB 490 D-EGVAPAAQIGELFAALILVPRDQALRELEEVLLKALAINSVHIDEPRIMSS 448
 QY 470 SEPTAHINAVDTAV-VWIAKSGISSGQSMPIANIMLSHVPHASNKCHELLNMCAN 428
 DB 449 CKRIH-----PALLIFRAGAGKQSCAPRRACSHITIDPIIQTAKVIAHRYVKI 502
 QY 429 KNAVYVDLLLEMINA 444
 DB 503 ECKVPMKLEFLEMA 518

RESULT 17
 A39596
 Progesterone receptor B form - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 20-Aug-1999
 A:Accession: A39596; I49111
 R:Shiota, E.R.; Shymala, G.; Schneider, W.; Parry, G.
 Biochemistry 30, 7014-7020, 1991
 A:Title: Molecular cloning, sequence analysis, and expression of complementary DNA en
 A:Reference number: A39596; MIM: 91299759; PMID: 2069958
 A:Accession: A39596
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-923 <SCH>
 A:Cross-references: GB: M68915, GB: J05333, NID: q200471, PDB: AAA9972.1, PDB: q200472
 R:Bhatnagar, K.; Wu-Peng, X.S.; Fushasbi, T.; Kato, J.; Platt, D.W.
 Biochem. Biophys. Res. Commun. 205, 1093-1101, 1994
 A:Title: Nucleic acid sequence and DNAse hypersensitivity sites of the 5' region of the
 A:Reference number: I49111; MIM: q4910043; PMID: 7602647
 A:Accession: I49111
 A:Status: preliminary; translated from GB/EMBL/DDHU
 A:Molecule type: DNA
 A:Residues: 1-9 SHAC
 A:Cross-references: EMBL: U12644; NID: q639916; PDB: AAA66067.1, PDB: q639917
 C:Superfamily: progesterone receptor; erba transforming protein homology
 C:Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation;
 F:557-577/Region; erba transforming protein homology <ERBA>
 F:593-617/Region; zinc finger

Query Match 18.58; Score 466.5; DB 2; Length 923;
 Best local similarity 30.98; Pred. No. 5, 80-29;
 Matches 114; Conservative 65; Mismatches 115; Indels 75; Gaps 13

QY 96 CAACSDYASGTHGVASCCEGCAKAFKRSIQGHNIYICATNCTIDKAPKSCAGRLRK 155
 DB 557 CACVCHASGVHGVASCPAKAFKRTIDGNIYVCPATINEFTIKPKKSCAGCPKPV 616
 QY 156 CTEVGVWVKQNSP--PPGCVPLVPPSPADE---GLHCAKAKKSGCHADVPPELILA--SPE 213

100 617 GCGAGVAGLGGKPKK - FNVVVRALD----- -CVALLQSVAPNESQTLGGRIIPSN 668
 214 G-----LVITLLEAPPIVLLSRISA -PFTASMMMSITKIAKELVHMSMAKRTTG 265
 669 QETGVVPLINLMSTIPDVVAGHNTKIDTSSLSLSTNOIAKQGLASVYKWSKSLDG 728
 266 FVELSTLQVRLLESTWMEVIMMGLMWRSLDING - KLIAPPLVLDRECKVRELTLEI 423
 729 FRNLHIDQVLTLLQVSSMSLMVGTLMKRSYKRVSGQMLTFATPLLINQRMK - ELISF 784
 424 FIMLLA -TTSPPRELKIQKRYLVRKAMILLNSSMYPLVATQVAFSSKRLAHILNAV 480
 785 YSLTCTVMQVLPQEFVKIQVTHREPLCMKVLILLNT - - - - - ELIRSSQFEMRSY 848
 481 TDALVWVIAKSGISQGSMSRLANILMSHPVLSANRKNRDEPLINRCKNVVY - YD 446
 820 - - - - - ILEGLRSGSQQFEMRS - - - - - SVTPE - - - - - LKAKVLRGVVTFSSQRPV 862
 447 L LLEMIN 443
 863 LKRLDSIH 871

RESULT 18

153280
 progesterone receptor B form

C/Spectre: Kallus novotus (Morway rat)

C/Date: 02 Aug 1996 #sequence revision 02-Aug-1996 #next_change 20 Aug 1999

C/Accession: U53280; B24743; A49574; A23743

R/Date: Satou, O.K.; Mayo, K.E.

R/Endofunc: 134, 709-718, 1994

A/Title: Isolation of the progesterone receptor gene by subtractive cloning and cDNA library

A/Reference number: U53280; M01D:94140817; PMID:829566

A/Accession: U53280

A/Status: preliminary; translated from GR/EMBL/DBEM

A/Molecule type: mRNA

A/Cross-reference: GR:U53280; NID:3463282; PUD:AAA19916.1; PID:3463283

R/Date: O.K.; Mayo, K.E.

R/Endofunc: 5, 967-978, 1991

A/Title: Transient expression of progesterone receptor messenger RNA in ovarian granulosa

A/Reference number: A23743; M01D:92049379; PMID:1840636

A/Accession: B23743

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A/Molecule type: mRNA

A/Cross-reference: GR:U53280; W:591.624;742.910 - PWR2

R/Date: W.L.; Montano, M.M.; Katznelson, B.S.

R/Endofunc: 7, 1603-1616, 1993

A/Title: Cloning of the rat progesterone receptor gene 5' region and identification of B

A/Reference number: A49574; M01D:94195318; PMID:8145766

A/Accession: A49574

A/Status: preliminary; translated from GR/EMBL/DBEM

A/Molecule type: DNA

A/Cross-reference: GR:U53280; NID:9466178

C/Spectre: progesterone receptor cDNA transforming protein hemology

C/Keywords: RNA binding; nucleic acid binding; protein receptor; zinc finger

C/557 819/96/ma: cDNA transforming protein hemology - ERNA

C/557 577/96/ma: zinc finger

C/557 617/96/ma: zinc finger

C/557 617/96/ma: zinc finger

C/557 617/96/ma: zinc finger

C/557 617/96/ma: zinc finger

100 617 GCGAGVAGLGGKPKK - FNVVVRALD----- -CVALLQSVAPNESQTLGGRIIPSN 668
 214 G-----LVITLLEAPPIVLLSRISA -PFTASMMMSITKIAKELVHMSMAKRTTG 265
 669 QETGVVPLINLMSTIPDVVAGHNTKIDTSSLSLSTNOIAKQGLASVYKWSKSLDG 728
 266 FVELSTLQVRLLESTWMEVIMMGLMWRSLDING - KLIAPPLVLDRECKVRELTLEI 423
 729 FRNLHIDQVLTLLQVSSMSLMVGTLMKRSYKRVSGQMLTFATPLLINQRMK - ELISF 784
 424 FIMLLA -TTSPPRELKIQKRYLVRKAMILLNSSMYPLVATQVAFSSKRLAHILNAV 480
 785 YSLTCTVMQVLPQEFVKIQVTHREPLCMKVLILLNT - - - - - ELIRSSQFEMRSY 848
 481 TDALVWVIA -KSGISQGSMSRLANILMSHPVLSANRKNRDEPLINRCKNVVY - YD 446
 820 - - - - - ILEGLRSGSQQFEMRS - - - - - SVTPE - - - - - LKAKVLRGVVTFSSQRPV 862
 447 L LLEMIN 443
 863 LKRLDSIH 871

RESULT 19

84781
 retinoid X receptor beta - human

C/Spectre: Homo sapiens (man)

C/Date: 31-Dec-1993 #sequence revision 02-Aug-1994 #next_change 02-Aug-2002

C/Accession: S47781; S49753; S22647; A41727; E41727; A44270

R/Date: Schuster, K.; McBride, O.W.; DiSanto, J.P.; Ozato, K.; Yang, S.Y.

R/Endofunc: 90, 505-510, 1993

A/Title: Cloning and chromosome mapping of human retinoid X receptor beta: selective

A/Reference number: S47781; M01D:93144716; PMID:8481486

A/Accession: S47781

A/Status: preliminary

A/Molecule type: mRNA

A/Cross-reference: S49753

A/Status: preliminary

A/Molecule type: mRNA

A/Cross-reference: EMBL:X64522; NID:930447; PID:CA445087.1; PID:930444

R/Date: Schuster, K.; Park, J.H.; DiSanto, J.P.; Marks, M.; Ozato, K.; Yang, S.Y.

R/Endofunc: 20, 1801, 1992

A/Title: Isolation of a full-length cDNA clone encoding a N-terminally variant form of

A/Reference number: S22647; M01D:92253386; PMID:1315958

A/Accession: S22647

A/Status: preliminary

A/Molecule type: mRNA

A/Cross-reference: EMBL:X64522

R/Date: M.J.; Kastner, P.; Lyons, K.; Nakshatri, H.; Saunders, M.; Zacharewski, T.; Cho

R/Endofunc: 68, 477-495, 1992

A/Title: Purification, cloning, and RAR identity of the beta cell factor with which R

A/Reference number: A41727; M01D:92127595; PMID:1310259

A/Accession: A41727

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

A/Cross-reference: S49753; S22647; NID:930447; PID:CA445087.1; PID:930444

A/Status: preliminary

A/Molecule type: mRNA

A/Cross-reference: S49753; S22647; NID:930447; PID:CA445087.1; PID:930444

A/Status: preliminary

A/Molecule type: mRNA

A/Cross-reference: S49753; S22647; NID:930447; PID:CA445087.1; PID:930444

Citations:

Accession: GDB:RXRB
 Accession: references: 34p 128001; mw: 18,294
 A:Map position: 6p21.3-6p21.3
 C:Superfamily: retinoic acid receptor alpha; etra transforming protein homology
 Keywords: zinc finger
 F1201-449/bemman: etra transforming protein homology -EBBA-

Query Match 18.1% Score 458.5; 38 2; Length 533;
 Best Local Similarity 40.2%; Pred. No. 1,30-28;
 Matches 140; Conservative 63; Mismatches 163; Indels 97; Gaps 17;

QY 9 VTNIKPGGPGTISFVNIWTFPHSHIVVHFGSHLVAFKPSKWTAKSLHTLVN 67
 DB 136 ISSMSSTGPPAP-----PG-----ESGVSSPGINSTV---SLPGG 172
 QY 68 ---PRTKPKVGNR CASVITPGSKRIAHFAVGSVAVSGYVWSTFGKAFK 121
 DB 173 SGPPEVKKPVLIVPGIHHPGPPGPAK--RIVATGQSSCKHYGVSTGPKGPK 230
 QY 122 PSTGQINIVTGTATNCTTTRKNPKSCVAPLRYGVWVKGCGSPPEKIVPVPGPS 181
 DB 231 PTLPKRI TYSQGNKPTVTKRPNQGVYGVGKLTATMKRPVAVPF-----PGGG 159
 QY 182 APEQIHVAVKPSGCHAVPVPTITPAISPGQ VITLIPAPPHVIT 229
 DB 283 KKPRT --GPGDAGG-----APEEMVDSLEFAT--LAVFYSKYVGGP 199
 QY 230 -----SPGAPPTASMMKSTTKLAPETVIMISWAKTPGVPVSLFVVPITPGW 282
 DB 123 GATGSGSSPNQVVT-----NIGVADQGLPTIVWAKPTIHPSILPDQVITLPAKW 254
 QY 283 MYLMMGIMPSIDHDKTIFADPVLDPKCVGPIIFIDPMILA-TTSFPTKIQH 341
 DB 377 NEILLASPSRPTVPGVITLAIQVHVRSAISA GVATITVPTITVSMRPMRMR 412
 QY 342 KEYLCVAKMLLNSMRYPIVATQDSSPKLAHLINAVTALVWVIAKSGTSSQSGSR 401
 DB 436 TPLGCPALITLFPDPAKGL-----SNPSEVPIPEKVASLFTYGVKVPYGG--P 485
 QY 402 LANLMLSHVHASNPKMEHLNMCKKNVVPYLLLEMLNA 444
 DB 486 FAKLLRLPALRSLGKLEHLEFFKLTGPTIDPLMLMLBA 528

RESULT 20

S26669

Retinoic acid receptor X-beta - mouse

N:Alternate names: RXR-beta2 isoform

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999

C:Accession: S26669; 163160

R:Model:SDOT; D:0.3; homology: 0; Heyman, R.A.; Zhou, J.Y.; Ong, E.S.; Oro, A.E.; Kaki

Genes Dev. 6, 329-344, 1992

A:Title: Characterization of three RXR genes that mediate the action of 9-cis retinoic

A:Reference number: S26668; MIMD:9192447; PMID:113197

A:Accession: S26669

A>Status: preliminary

A:Molecule type: mRNA

A:Accession: 1-410 <MAN>

A:Cross-references: EMBL:X66224; NID:945024; PID:GAA6962.1; PID:945024

R:Humada, K.; Kono, Y.; Ozato, K.; Taketo, M.

Gene 142, 183-189, 1994

A:Title: The mouse RXRb gene encoding RXR beta: genomic organization and two mRNA isoform

A:Reference number: 148752; MIMD:9422555; PMID:8194750

A:Accession: 163160

A>Status: preliminary; translated from GR/EMBL/DBM1

A:Molecule type: DNA

A:Accession: 1-410 <PFS>

A:Cross-references: GR:D21831; NID:9497828; PID:BA04859.1; PID:9497828

C:Species: RXRB

A:Introns: 38/7; 9/71 15/71 27/71 29/71 37/71 39/71

C:Superfamily: retinoic acid receptor alpha; etra transforming protein homology
 Keywords: zinc finger
 E:50 45/Database: etra transforming protein homology -EBBA-

Query Match 18.1% Score 455.5; 38 2; Length 410;
 Best Local Similarity 40.2%; Pred. No. 1,60-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

QY 9 VTNIKPGGPGTISFVNIWTFPHSHIVVHFGSHLVAFKPSKWTAKSLHTLVN 67
 DB 13 ISSMSSTGPPAP-----PG-----ESGVSSPGINSTV---SLPGG 49
 QY 68 ---PRTKPKVGNR CASVITPGSKRIAHFAVGSVAVSGYVWSTFGKAFK 121
 DB 50 SGPPEVKKPVLIVPGIHHPGPPGPAK--RIVATGQSSCKHYGVSTGPKGPK 107
 QY 122 PSTGQINIVTGTATNCTTTRKNPKSCVAPLRYGVWVKGCGSPPEKIVPVPGPS 181
 DB 108 PTLPKRI TYSQGNKPTVTKRPNQGVYGVGKLTATMKRPVAVPF-----PGGG 159
 QY 182 APEQIHVAVKPSGCHAVPVPTITPAISPGQ VITLIPAPPHVIT 229
 DB 160 KKPRT --GPGDAGG-----APEEMVDSLEFAT--LAVFYSKYVGGP 199
 QY 230 -----SPGAPPTASMMKSTTKLAPETVIMISWAKTPGVPVSLFVVPITPGW 282
 DB 200 GATGSGSSPNQVVT-----NIGVADQGLPTIVWAKPTIHPSILPDQVITLPAKW 254
 QY 283 MYLMMGIMPSIDHDKTIFADPVLDPKCVGPIIFIDPMILA-TTSFPTKIQH 341
 DB 354 NEILLASPSRPTVPGVITLAIQVHVRSAISA GVATITVPTITVSMRPMRMR 412
 QY 342 KEYLCVAKMLLNSMRYPIVATQDSSPKLAHLINAVTALVWVIAKSGTSSQSGSR 401
 DB 412 TPLGCPALITLFPDPAKGL-----SNPSEVPIPEKVASLFTYGVKVPYGG--P 462
 QY 402 LANLMLSHVHASNPKMEHLNMCKKNVVPYLLLEMLNA 444
 DB 463 FAKLLRLPALRSLGKLEHLEFFKLTGPTIDPLMLMLBA 405

RESULT 21

A34418

H-2 region II binding protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 15-Jun-1992 #sequence_revision 09-Oct-1992 #text_change 02-Aug-2002

C:Accession: A34418

R:Humada, K.; Gleason, S.L.; Levy, B.Z.; Hirschfeld, S.; Appella, E.; Ozato, K.

Proc. Natl. Acad. Sci. U.S.A. 86, 8289-8294, 1989

A:Title: H 2RIIRP, a member of the nuclear hormone receptor superfamily that binds to

A:Reference number: A34418; MIMD:90046782; PMID:2554107

A:Accession: A34418

A>Status: preliminary

A:Molecule type: mRNA

A:Accession: 1-446 <HMS>

A:Cross-references: GR:M26834; NID:9193740; PID:AAA5772.1; PID:9487184

C:Superfamily: retinoic acid receptor alpha; etra transforming protein homology

Keywords: DNA binding; transcription regulation; zinc finger

F116 362/bemman: etra transforming protein homology -EBBA-

Query Match 18.1% Score 455.5; 38 2; Length 446;
 Best Local Similarity 40.2%; Pred. No. 1,80-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

QY 9 VTNIKPGGPGTISFVNIWTFPHSHIVVHFGSHLVAFKPSKWTAKSLHTLVN 67
 DB 49 ISSMSSTGPPAP-----PG-----ESGVSSPGINSTV---SLPGG 85
 QY 68 ---PRTKPKVGNR CASVITPGSKRIAHFAVGSVAVSGYVWSTFGKAFK 121
 DB 86 SGPPEVKKPVLIVPGIHHPGPPGPAK--RIVATGQSSCKHYGVSTGPKGPK 144
 QY 122 PSTGQINIVTGTATNCTTTRKNPKSCVAPLRYGVWVKGCGSPPEKIVPVPGPS 181

```

144 KTLKRLTYSKRNK*TVKRGKNGYVTFVGLATGMRKEAVGFT
182 ADPGLRACKAKKSGSHAVRKLALSTPGL VTIILAEPRVILL
196 KDKD GGGGAGG-----APEDMVDRIILAE--LAVESKSLGGVDP 245
240 -----SRPSAFETASMMKSTLKLADKELVIMISWAKIPGVNLSLPGVRLNSGW 282
246 GATGGGSSNDPVT-----NIGCAADKQLFTIVKAKRIPHSLSLIDPVILLAKW 289
284 MEVIMGLMRSSTIDHKKLFAVDVLDDEKCVGSLLEFDMLA-TSPREKLQH 441
290 NEILLASFHRSIDVHMLTLATGLHVNHSASHA-GVCAITLAVLTETVSKMDEMEEK 446
442 KEIVAKAMLLNSSWYPIVATVQDSSKRLAHLLNAVTDALVWVAKSSISQGSMSR 401
449 TELACTALILFMDAKGL-----SNRGVEILREKVVASLFTYKQKVPDQGG-R 498
402 IANIMLSIVHVASNKGMHLLNMCKNVAVVYDILDEMLNA 444
499 FAKLLRLPALPSIGKLEHLPFKLIDITIDPELMMLFA 441

```

RESULT 22

18.18: Score 455.5; DB 2; Length 448;
 best local similarity 40.28; Pred. No. 1.8e-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

Query Match
 18.18: Score 455.5; DB 2; Length 448;
 best local similarity 40.28; Pred. No. 1.8e-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

9 VTINEDDTPGKOTISPVNIWPTDHSPIVNHQSLSHLVAEYKSWCFARSLHLLPVN 67
 51 ISSMSKSPGLPPAP-----PSG-----PSGVSSDGLNSTV--SLDAG 87
 68 KETLKRVSNP--CASVYDQSKRIMHPVAVSNVASYHYVWMSFGSKAFKK 121
 88 SGPENKPVVLEKRLHLPVPGDQAK--KLVADQKSSAHNVYVSTHGRKGRK 145
 122 KSTQGNIDYGVATNGTIDKRNKNSGVAKLKYUYEVGVKSGSRKEGCVLRVROPS 181
 146 KTLKRLTYSKRNK*TVKRGKNGYVTFVGLATGMRKEAVGFT-----RQGG 197
 182 ADPGLRACKAKKSGSHAVRKLALSTPGL VTIILAEPRVILL----- 229
 196 KDKD GGGGAGG-----APEDMVDRIILAE--LAVESKSLGGVDP 247
 240 -----SRPSAFETASMMKSTLKLADKELVIMISWAKIPGVNLSLPGVRLNSGW 282
 246 GATGGGSSNDPVT-----NIGCAADKQLFTIVKAKRIPHSLSLIDPVILLAKW 289
 284 MEVIMGLMRSSTIDHKKLFAVDVLDDEKCVGSLLEFDMLA-TSPREKLQH 441
 290 NEILLASFHRSIDVHMLTLATGLHVNHSASHA-GVCAITLAVLTETVSKMDEMEEK 446
 442 KEIVAKAMLLNSSWYPIVATVQDSSKRLAHLLNAVTDALVWVAKSSISQGSMSR 401
 449 TELACTALILFMDAKGL-----SNRGVEILREKVVASLFTYKQKVPDQGG-R 498
 402 IANIMLSIVHVASNKGMHLLNMCKNVAVVYDILDEMLNA 444
 499 FAKLLRLPALPSIGKLEHLPFKLIDITIDPELMMLFA 441

Query Match
 18.18: Score 455.5; DB 2; Length 448;
 best local similarity 40.28; Pred. No. 1.8e-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

9 VTINEDDTPGKOTISPVNIWPTDHSPIVNHQSLSHLVAEYKSWCFARSLHLLPVN 67
 51 ISSMSKSPGLPPAP-----PSG-----PSGVSSDGLNSTV--SLDAG 87
 68 KETLKRVSNP--CASVYDQSKRIMHPVAVSNVASYHYVWMSFGSKAFKK 121
 88 SGPENKPVVLEKRLHLPVPGDQAK--KLVADQKSSAHNVYVSTHGRKGRK 145
 122 KSTQGNIDYGVATNGTIDKRNKNSGVAKLKYUYEVGVKSGSRKEGCVLRVROPS 181
 146 KTLKRLTYSKRNK*TVKRGKNGYVTFVGLATGMRKEAVGFT-----RQGG 197
 182 ADPGLRACKAKKSGSHAVRKLALSTPGL VTIILAEPRVILL----- 229
 196 KDKD GGGGAGG-----APEDMVDRIILAE--LAVESKSLGGVDP 247
 240 -----SRPSAFETASMMKSTLKLADKELVIMISWAKIPGVNLSLPGVRLNSGW 282
 246 GATGGGSSNDPVT-----NIGCAADKQLFTIVKAKRIPHSLSLIDPVILLAKW 289
 284 MEVIMGLMRSSTIDHKKLFAVDVLDDEKCVGSLLEFDMLA-TSPREKLQH 441
 290 NEILLASFHRSIDVHMLTLATGLHVNHSASHA-GVCAITLAVLTETVSKMDEMEEK 446
 442 KEIVAKAMLLNSSWYPIVATVQDSSKRLAHLLNAVTDALVWVAKSSISQGSMSR 401
 449 TELACTALILFMDAKGL-----SNRGVEILREKVVASLFTYKQKVPDQGG-R 498
 402 IANIMLSIVHVASNKGMHLLNMCKNVAVVYDILDEMLNA 444
 499 FAKLLRLPALPSIGKLEHLPFKLIDITIDPELMMLFA 441

```

442 KEIVAKAMLLNSSWYPIVATVQDSSKRLAHLLNAVTDALVWVAKSSISQGSMSR 401
449 TELACTALILFMDAKGL-----SNRGVEILREKVVASLFTYKQKVPDQGG-R 498
402 IANIMLSIVHVASNKGMHLLNMCKNVAVVYDILDEMLNA 444
499 FAKLLRLPALPSIGKLEHLPFKLIDITIDPELMMLFA 441

```

RESULT 23

18.18: Score 455.5; DB 2; Length 451;
 best local similarity 40.28; Pred. No. 1.8e-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

Query Match
 18.18: Score 455.5; DB 2; Length 451;
 best local similarity 40.28; Pred. No. 1.8e-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

9 VTINEDDTPGKOTISPVNIWPTDHSPIVNHQSLSHLVAEYKSWCFARSLHLLPVN 67
 51 ISSMSKSPGLPPAP-----PSG-----PSGVSSDGLNSTV--SLDAG 87
 68 KETLKRVSNP--CASVYDQSKRIMHPVAVSNVASYHYVWMSFGSKAFKK 121
 91 SGPENKPVVLEKRLHLPVPGDQAK--KLVADQKSSAHNVYVSTHGRKGRK 148
 122 KSTQGNIDYGVATNGTIDKRNKNSGVAKLKYUYEVGVKSGSRKEGCVLRVROPS 181
 149 KTLKRLTYSKRNK*TVKRGKNGYVTFVGLATGMRKEAVGFT-----RQGG 200
 182 ADPGLRACKAKKSGSHAVRKLALSTPGL VTIILAEPRVILL----- 229
 201 KDKD GGGGAGG-----APEDMVDRIILAE--LAVESKSLGGVDP 240
 240 -----SRPSAFETASMMKSTLKLADKELVIMISWAKIPGVNLSLPGVRLNSGW 282
 241 GATGGGSSNDPVT-----NIGCAADKQLFTIVKAKRIPHSLSLIDPVILLAKW 294
 284 MEVIMGLMRSSTIDHKKLFAVDVLDDEKCVGSLLEFDMLA-TSPREKLQH 441
 295 NEILLASFHRSIDVHMLTLATGLHVNHSASHA-GVCAITLAVLTETVSKMDEMEEK 446
 442 KEIVAKAMLLNSSWYPIVATVQDSSKRLAHLLNAVTDALVWVAKSSISQGSMSR 401
 454 TELACTALILFMDAKGL-----SNRGVEILREKVVASLFTYKQKVPDQGG-R 404
 402 IANIMLSIVHVASNKGMHLLNMCKNVAVVYDILDEMLNA 444
 404 FAKLLRLPALPSIGKLEHLPFKLIDITIDPELMMLFA 446

Query Match
 18.18: Score 455.5; DB 2; Length 451;
 best local similarity 40.28; Pred. No. 1.8e-28;
 Matches 140; Conservative 62; Mismatches 164; Indels 97; Gaps 17;

9 VTINEDDTPGKOTISPVNIWPTDHSPIVNHQSLSHLVAEYKSWCFARSLHLLPVN 67
 51 ISSMSKSPGLPPAP-----PSG-----PSGVSSDGLNSTV--SLDAG 87
 68 KETLKRVSNP--CASVYDQSKRIMHPVAVSNVASYHYVWMSFGSKAFKK 121
 91 SGPENKPVVLEKRLHLPVPGDQAK--KLVADQKSSAHNVYVSTHGRKGRK 148
 122 KSTQGNIDYGVATNGTIDKRNKNSGVAKLKYUYEVGVKSGSRKEGCVLRVROPS 181
 149 KTLKRLTYSKRNK*TVKRGKNGYVTFVGLATGMRKEAVGFT-----RQGG 200
 182 ADPGLRACKAKKSGSHAVRKLALSTPGL VTIILAEPRVILL----- 229
 201 KDKD GGGGAGG-----APEDMVDRIILAE--LAVESKSLGGVDP 240
 240 -----SRPSAFETASMMKSTLKLADKELVIMISWAKIPGVNLSLPGVRLNSGW 282
 241 GATGGGSSNDPVT-----NIGCAADKQLFTIVKAKRIPHSLSLIDPVILLAKW 294
 284 MEVIMGLMRSSTIDHKKLFAVDVLDDEKCVGSLLEFDMLA-TSPREKLQH 441
 295 NEILLASFHRSIDVHMLTLATGLHVNHSASHA-GVCAITLAVLTETVSKMDEMEEK 446
 442 KEIVAKAMLLNSSWYPIVATVQDSSKRLAHLLNAVTDALVWVAKSSISQGSMSR 401
 454 TELACTALILFMDAKGL-----SNRGVEILREKVVASLFTYKQKVPDQGG-R 404
 402 IANIMLSIVHVASNKGMHLLNMCKNVAVVYDILDEMLNA 444
 404 FAKLLRLPALPSIGKLEHLPFKLIDITIDPELMMLFA 446

Query Match	18.0%	Score 453	Pr 2	Length 400
Best Local Similarity	31.68%			
			Pred No. 70-28	

[illegible]

Superfamily: retinoid acid receptor alpha; erba transforming protein homolog
 Keywords: DNA binding; transcription regulation; zinc finger
 P118-183 domain: erba transforming protein homology - ERPA

Query Match 17.4% Score 437.5 DB 2: Length 467
 Best Local Similarity 28.0% Pos No 3 40-27
 Matches 140: Conservative 68; Mismatches 205; Indels 72; Gaps 14

UY 1 MNYSLPSNVNLEGGPQTTSPNVIWPT--PRHISPT---VHPPQISHVA--EPKPS 52
 DB 5 LDESLGVNSSSSISPPGSGMAAPSLHSLSPGLSPGLSPGLSPHSLSPINSGP 68
 OY 54 PMVEARST--PHITPV--NPFTIKPVYSCNPKASPVGCGSKDP----- 90
 DB 69 PPSVSISSPMGPHISMVPTTPTLDFEGSPQLNSPMNPVSSSDIKPRLGLNGVKVPAHP 128
 OY 93 -----AMPAVSTVASVYHVAWSEETKAFKPSLPHNLYTPATNWTPLVNP 144
 DB 129 SCNMSSFTKHITACIDPSSSKHYVYSCPKCKPKPTKPTITPTDNPVPTLPKQ 188
 OY 145 PKSCGCPLEKFEYVMVCGSPRPCTPLVPPGSKAPQVTAAPASCTAAFPVPT 204
 DB 189 KNPVYTPYKQTLAMKMPFAVDE-----KQKIDKNENVESTSSAMEMP--VEK 239
 OY 205 ILIDALSPFALVITLFA---TPPHVLTSPSAPTEVSMMSTTVAPEVHVTSMA 260
 DB 240 ILPAHLAVEIKETLEYEAMKGLNP-----SSNPQVY-----NIOADKQLFLVEMA 288
 OY 264 KPIHPEVPICTITVNVLEETVWPFVMMHMMPS--THPEKLTATLVLEDEEYGVN 320
 DB 289 KPIHPEVPICTITVNVLEETVWPFVMMHMMPS--THPEKLTATLVLEDEEYGVN 347
 OY 321 TTECEMIA--TSEFEKCTQHPFN--YVAMKLTNCGMTVTATGCA--SSEFLMHLRA 379
 DB 348 GALTIVLTSLYSKMKMDMDKDTLGLCLATVLEFNDSGLSNPALVEALREKVVASLEA 407
 OY 380 VTDAIWWVIAKSGISQDSQSMRIANLMLSHVFRASNGMPEHLMNCKNVVYETLL 439
 DB 408 YCKH-----KRPQDGRPAKLTLPALRSIGIKPIHLEFPFKIGTPTDITLM 457
 OY 440 EMINA 444
 DB 458 EMLEA 462

RESULT 34
 A39248
 androgen receptor - human
 C1Species: Homo sapiens (man)
 C1Date: 04-Oct-1991 #sequence revision 04-Oct-1991 #text change 24-Nov-1999
 C1Accession: A39248; A0428; A40109; A6994; A4942; A2753; A1019; A1494; A4224; A13
 K1ubahn, D.B.; Brown, T.R.; Simmler, J.A.; Higgs, H.N.; Miquen, C.J.; Wilson, E.M.; Pe
 Proc. Natl. Acad. Sci. U.S.A. 86, 9544-9548, 1989
 A1Title: Sequence of the intron/exon junctions of the coding region of the human androg
 A1Reference number: A39248; MIM:190833;2; PMID:2594783
 A1Accession: A39248
 A1Molecule type: DNA
 A1Residues: 1-919 <10B>
 A1Status: not compared with conceptual translation
 A1Cross-references: GR:M2743; GR:M2743; MIM:317904; PDB:AAE1774; PDB:AAE1774; PDB:AAE1774
 K1ubahn, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, D.F.; French, F.S.; Wilson, E.M.
 Mol. Cell. Endocrinol. 61, 257-262, 1989
 A1Title: The N-terminal domain of the human androgen receptor is encoded by one, large c
 A1Reference number: A30428; MIM:8911730; PMID:2917688
 A1Accession: A30428
 A1Molecule type: DNA
 A1Residues: 1-77,79-165,74,167-389,71,391-464,474-538 <EPA>
 A1Cross-references: GB:M20260
 K1ubahn, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, D.F.; French, F.S.; Wilson, E.M.
 Science 240, 327-330, 1988
 A1Title: Cloning of human androgen receptor complementary DNA and localization to the X
 A1Reference number: A40109; MIM:8817811;2; PMID:3353727
 A1Accession: A40109
 A1Molecule type: DNA

A1Residues: 559-624 <10D>
 A1Cross-references: GB:M20132
 K1ubahn, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, D.F.; French, F.S.; Wilson, E.M.
 J. Mol. Endocrinol. 2, 81-84, 1989
 A1Title: Structural organization of the human androgen receptor gene.
 A1Reference number: A60946; MIM:89322749; PMID:2546571
 A1Accession: A60946
 A1Molecule type: DNA
 A1Residues: 546-640,587-604,626-631,723-726,770-774,814-818,867-870 <K10>
 K1ubahn, D.B.; Joseph, D.R.; Sar, M.; Tan, J.; Higgs, H.N.; Larson, R.E.; French, F.
 Mol. Endocrinol. 2, 1265-1275, 1988
 A1Title: The human androgen receptor: complementary deoxyribonucleic acid cloning, se
 A1Reference number: A39442; MIM:89112208; PMID:3216866
 A1Accession: A39442
 A1Molecule type: mRNA
 A1Residues: 1-919 <103>
 A1Cross-references: GB:M29132; MIM:3178627; PDB:AAE1729; PDB:AAE1729; PDB:AAE1729; PDB:AAE1729
 R1Trapani, J.; Klaassen, P.; Kuiper, G.G.J.M.; van der Korf, J.A.G.M.; Faber, L.W.
 Blooom, Biophys. Res. Commun. 153, 241-248, 1988
 A1Title: Cloning, structure and expression of a cDNA encoding the human androgen rece
 A1Reference number: A27653; MIM:88440407; PMID:3377788
 A1Accession: A27653
 A1Molecule type: mRNA
 A1Residues: 468-564,7K,566-919 <TRA>
 A1Cross-references: GB:M29260; MIM:3178891; PDB:AAE1774; PDB:AAE17892
 A1Note: the authors translated the codon AAG for residue 565 as glu
 Reichard, C.; Kokontis, J.; Liao, S.
 Science 240, 324-326, 1988
 A1Title: Molecular cloning of human and rat androgen receptor
 A1Reference number: A40108; MIM:8817811;2; PMID:3354726
 A1Accession: A40108
 A1Molecule type: mRNA
 A1Residues: 557-628 <10D>
 A1Cross-references: GB:M18624
 Reichard, C.; Kokontis, J.; Liao, S.
 Proc. Natl. Acad. Sci. U.S.A. 85, 7211-7215, 1988
 A1Title: Structural analysis of complementary DNA and amino acid sequences of human a
 A1Reference number: A40944; MIM:89017168; PMID:3174628
 A1Accession: A40944
 A1Molecule type: mRNA
 A1Residues: 1-74,79-89,71,90-472,506,473-474,476-644,7K,646-919 <CH2>
 A1Cross-references: GB:M24263
 R1Tilley, W.D.; Marcolini, M.; Wilson, J.D.; McPhaul, M.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 327-331, 1989
 A1Title: Characterization and expression of a cDNA encoding the human androgen recept
 A1Reference number: A32224; MIM:89089999; PMID:2911578
 A1Accession: A32224
 A1Molecule type: mRNA
 A1Residues: 1-77,79-211,7K,213-471,473-919 <11D>
 A1Cross-references: GB:M21748; GR:M204150; MIM:3178871; PDB:AAE1771; PDB:AAE17892
 E1Marsden, J.; Lee, H.J.; Cheng, H.T.; Mestayer, C.; Porcillo, M.C.; Gabboli, S.; Mar
 Mol. Endocrinol. 7, 861-869, 1993
 A1Title: A point mutation in the second zinc finger of the DNA-binding domain of the
 A1Reference number: A40125; MIM:94019395; PMID:8413410
 A1Accession: A40125
 A1Status: not compared with conceptual translation
 A1Molecule type: DNA
 A1Residues: 557-614,617-619,624-624 <M4>
 A1Cross-references: PDB:AAH28340; PDB:AAH28340
 C1Accession: A39248
 A1Cross-references: GB:M20556; MIM:313700
 A1Map position: X111-X912
 A1Residues: 538-599,601-604,606-607,609-610,612-613
 C1Superfamily: classed erba related proteins; erba transforming protein homolog
 C1Keywords: DNA binding; steroid binding; transcription regulation; zinc finger
 E1559-579/Region: zinc finger
 E1595-619/Region: zinc finger

Query Match 17.4% Score 437.5 DB 2: Length 919
 Best Local Similarity 26.8% Pos No 1 20-26

Keywords: zinc finger
 F548-R66/Domain: etla transforming protein homology <ERBA>
 F550-570/Region: zinc finger
 F586-610/Region: zinc finger

Query Match 17.28, Score 434, DB 2, Length 911:
 Best Local Similarity 28.64, Prod No 2.26-26;
 Matches 109; Conservative 72; Mismatches 108; Idels 92; Gaps 14;

```

QY      1  EGGGKROT--TSPNVLP-----TPGH,SPLVHRQLSHLYAFDQKSHPCARS----- 59
      2  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      478  GGIAGGFSQFTAPVWVPGGVMSVPD-YDPTGVKSEM-----GPMMDSYGQPYD 527
      3  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      60  -----LEHTLPVNRKELFKVSCNPGASPVTPGSGKPVAMH-----CAVNSQVANSQTHY 108
      4  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      528  MRLETDADHVLPI-----DYPPPOKTKTLGGDEASGCHY 562
      5  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      109  GVNSTCEGKAFKFKSLGHNVTCPATNQTTLENFESQACZPLPTLYEVMVNGSR 168
      6  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      563  GALTCTGSCVFFKRAFGQGYTCASPNVTTLTKPPKNCPSGLPKVEACMT LGAP- 420
      7  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      169  ERQAVPIVPPQSAIDQIHCAQKAKRSGHAPVRELL--LDALSPQGLVILLAEPP 224
      8  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      621  -----KLRKLINLKIQEEHVSASTTSPTFETTKLIVSHIEGYTPPLNVIAIE 672
      9  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      225  PHYLIS--RPSAFTTASMMMSITKLADKELVIMISMAKKITGFVFLSLFDVKKLLRSGW 282
      10 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      673  PGVVGCHGNQNNQDPSFAIISINIGFQIVLVVWAKAPDLRNHIVNDQMAVLYQSW 742
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      283  MEVIMMGLMWRSLDHHK--LIFAPDLVLD-----RDEGKVGEGILIEFDMLLATISR 433
      12 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      743  MGLMVFAMKMRSLFTNNSRMLYFAPDLVFNENYEMKSRMYSQV-----MRHLSQE 784
      13 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      434  FRELKLGKREYLCVKAMITLN 454
      14 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      785  FGWLQITPDELICRMALLFS 805

```

Search completed: June 16, 2003, 14:16:36
 Job time : 28.4814 secs


```

DR EMBL: AF074598; AAC25602.1;
DR EMBL: AF074599; AAC25603.1;
DR EMBL: AF124790; AAC32580.1;
DR EMBL: AF047463; AAC3786.1;
DR EMBL: AF191544; AAC24232.1;
DR TRANSFAC: F04651;
DR Gene: HGNC:1468; ESR2.
DR MIM: 601663;
DR InterPro: IP000546; Hormone_rec_11q.
DR InterPro: IP001723; Steroid_receptor.
DR InterPro: IP001628; Zn1_Casteroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; z1-c4_1.
DR PRINTS: PR00498; STEROIDOMER.
DR PRINTS: PR00047; STEROIDOMER.
DR Prodom: PD00039; Zn1_Casteroid_1.
DR SMART: SM00430; HO11_1.
DR SMART: SM00499; Zn1_C4_1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation.
FT 1 MAIN 1 148 MODULATING.
FT 1 DNA_BIND 149 214 NUCLEAR RECEPTOR TYPE.
FT 1 ZN_FING 149 169 C4-TYPE.
FT 1 ZN_FING 185 209 C4-TYPE.
FT 1 DOMAIN 215 530 STEROID-BINDING.
FT 1 MOD_RES 87 87 PHOSPHORYLATION (BY SIMILARITY).
FT 1 MOD_RES 488 488 PHOSPHORYLATION (BY SIMILARITY).
FT 1 VARSPLIT 419 423 EYPLS -> WQNA (IN ISOFORM BETA-2A).
FT 1 VARSPLIT 319 409 MISSING (IN ISOFORM BETA-5A).
FT 1 VARSPLIT 324 530 MISSING (IN ISOFORM BETA-2A).
FT 1 VARSPLIT 365 375 DECKVGVILF -> YVPSGHDPG (IN ISOFORM BETA-6).
FT 1 VARSPLIT 376 530 MISSING (IN ISOFORM BETA-6).
FT 1 VARSPLIT 469 530 SNGGHEILNMKKNVAVYDILLEMNAHLVPRKSSITG
SESPAFDSKSKSGSNQSQ -> PAEASQTLISPMK
ELLPEALMEQ (IN ISOFORM BETA-2/CX)
SNGGHEILNMKKNVAVYDILLEMNAHLVPRKSSITG
SESPAFDSKSKSGSNQSQ -> NSLSKGLMPLHAS
CHGVRTGGAHMSVRSKSTFACQDRE (IN ISOFORM BETA-3).
SNGGHEILNMKKNVAVYDILLEMNAHLVPRKSSITG
SESPAFDSKSKSGSNQSQ -> KMEKQPIHLKLS
(IN ISOFORM BETA-4).
SNGGHEILNMKKNVAVYDILLEMNAHLVPRKSSITG
SESPAFDSKSKSGSNQSQ -> KVAL (IN ISOFORM BETA-5).
FT 1 CONFLICT 48 48 F > G (IN REF. ?).
FT 1 SPOONPCR 530 AA: 59216 MW: 804634215992418A CRC64.
Query Match 100.0% Score 2522; DR 1; Length 530;
Host Local Similarity 100.0%; Pred. No. 26-200;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

UY 401 LLEADLVLPDRGCKVVESSILPEPMILATISPEFKIKQREYVYKAMILLNSMYPL 360
DB 454 LLEADLVLPDRGCKVVESSILPEPMILATISPEFKIKQREYVYKAMILLNSMYPL 413
UY 461 VIALGVADSSPKLAMLNANVIALVWIAKNSLSSQGSMLANLMLLSIVRIASNGM 420
DB 414 VIALGVADSSPKLAMLNANVIALVWIAKNSLSSQGSMLANLMLLSIVRIASNGM 473
UY 421 EHLINMKKNVAVYDILLEMNAHLVPRKSSITGSESPAFDSKSGSNQSQ 477
DB 474 EHLINMKKNVAVYDILLEMNAHLVPRKSSITGSESPAFDSKSGSNQSQ 530
RESULT 2
ESR2_CALCIA
ID ESR2_CALCIA STANDARD; PRT; 530 AA.
AC 09517;
DT 16-OCT-2001 (Ref. 40, Cited)
DT 16-OCT-2001 (Ref. 40, Last sequence update)
DT 16-OCT-2001 (Ref. 40, Last annotation update)
DE Estrogen receptor beta (ER-beta).
GN ESR2 OR NR3A2.
OS Gallitrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Primatea; Platyrrhini; Callitrichidae.
OC Mammalia; Eulipotyria; Primates; Platyrrhini; Callitrichidae.
OC Gallitrix.
OC NCBI_TaxId:9484;
XP 1;
RP SEQUENCE FROM N.A.
RC L1000-TESTS;
PA Gaubhan J., Scobie G.;
BL Submitted (Nov 1999) to the PubMed/GenBank/TrEMBL databases.
CC -!- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
CC ER-ALPHA. AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
CC -!- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERO-DIMER WITH ER-
CC ALPHA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR3 SUBFAMILY.
CC
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.slb.ch/anonymous/
CC or send an email to license@isb.slb.ch).
CC
CC EMBL: Y09372; CA070546.2;
CC
CC HSSP: P03372; HCO.
CC InterPro: IP000536; Hormone_rec_11q.
CC InterPro: IP001723; Steroid_receptor.
CC InterPro: IP001628; Zn1_Casteroid.
CC Pfam: PF00104; hormone_rec_1.
CC Pfam: PF00105; z1-c4_1.
CC PRINTS: PR00498; STEROIDOMER.
CC PRINTS: PR00047; STEROIDOMER.
CC Prodom: PD00039; Zn1_Casteroid_1.
CC SMART: SM00430; HO11_1.
CC SMART: SM00499; Zn1_C4_1.
CC PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation.
FT 1 MAIN 1 148 MODULATING.
FT 1 DNA_BIND 149 214 NUCLEAR RECEPTOR TYPE.
FT 1 ZN_FING 149 169 C4-TYPE.
FT 1 ZN_FING 185 209 C4-TYPE.
FT 1 DOMAIN 215 530 STEROID-BINDING.
FT 1 MOD_RES 87 87 PHOSPHORYLATION (BY SIMILARITY).

```


DR	PROSTATE: P500041: NUCLEAR RECEPTOR: 1	
KM	Receptor: Transcription regulation: DNA-binding: Nuclear protein:	
KM	Zinc-finger: Steroid-binding: Alternative splicing: Glycoprotein:	
KM	Phosphorylation:	
ET	DOMAIN 1 148	MODULATING:
ET	DNA_BIND 149 214	NUCLEAR RECEPTOR-TYPE
ET	ZN_FING 149 169	CA-TYPE:
ET	ZN_FING 185 209	CA-TYPE:
ET	DOMAIN 215 530	STEROID-BINDING:
ET	MOD_RES 61 61	PHOSPHORYLATION (BY SIMILARITY):
ET	MOD_RES 87 87	PHOSPHORYLATION (BY MAJ): (PROBABLE):
ET	MOD_RES 105 105	PHOSPHORYLATION (BY MAJ): (PROBABLE):
ET	MOD_RES 488 488	PHOSPHORYLATION (BY SIMILARITY):
ET	CAROTIN 61 61	O-LINKED (GLICNA):
ET	VARSELIT 464 464	/FTTB-CAR-000201:
ET	VARSELIT 319 364	R -> S: ESSEPHRHWVAQKSAVPR (IN ISOFORM BETA-2):
ET	VARSELIT 365 409	MISSING (IN ISOFORM BETA-5):
ET	VARSELIT 319 409	MISSING (IN ISOFORM BETA-6):
ET	MUTAGEN 105 105	S -> A: ADELTHSES STIMULATORY EFFECT OF RAS:
ET	MUTAGEN 139 139	S -> A: NO LOSS OF THE STIMULATORY EFFECT OF RAS:
ET	CONFLICT 47 47	T -> A (IN REF. 2):
ET	CONFLICT 142 142	A -> T (IN REF. 2):
ET	CONFLICT 200 200	S -> R (IN REF. 2):
ET	CONFLICT 378 378	G -> T (IN REF. 2):
ET	CONFLICT 412 412	P -> R (IN REF. 2):
ET	CONFLICT 445 445	G -> R (IN REF. 2):
ET	CONFLICT 511 511	E -> G (IN REF. 2):
NO	SEQUENCE 530 AA: 59012 MW: 90E73A161550ABE4 CRF64:	
Query Match	89.28: Score 2250: DR 1: Length 530:	
Best Local Similarity	88.78: Pred. No. 5: 5e-178:	
Matches 424: Conservative 22: Mismatches 32: Indels 0: Gaps 0:		
UY	1 MNSVPSVNTLENSGKQTTSPNVIWPTPEISPLVWPGSHIYAFKSPWFAAS	66
DB	54 MNSVSSSTGNLEFGHVPKATASPNVIMPTSHISPLATGCSNITYAPQSPWFAAS	114
UY	61 EHTLPVNRFTKKKSGNCAAVTGTGSKDAHFCAMSDVASCYHGVMSDECKAF	120
DB	114 EHTLPVNRFTKKKSGNCAAVTGTGSKDAHFCAMSDVASCYHGVMSDECKAF	173
UY	121 KPSVGNHNYTPATNCTTPKPKPSQACPIKCTYVQKMGKSPSPKGYPLVPR	180
DB	174 KPSVGNHNYTPATNCTTPKPKPSQACPIKCTYVQKMGKSPSPKGYPLVPR	232
UY	181 SADECHCAKARPSGSHAPVPEIDAI SPQIVLTI FAPPPHVI ISPSAPETAS	240
DB	234 SADECHCAKARPSGSHAPVPEIDAI SPQIVLTI FAPPPHVI ISPSAPETAS	293
UY	241 MMSSTLFLATFTVHMSAKKIPGVEVETFGQVYVIFSGWMEVIMMCTWRETV	300
DB	294 MMSSTLFLATFTVHMSAKKIPGVEVETFGQVYVIFSGWMEVIMMCTWRETV	353
UY	301 LFAAPPIVLDPRGKCVETFTFPMGLATTSPPPEFKIQRHEVYGVAMITINS	360
DB	354 LFAAPPIVLDPRGKCVETFTFPMGLATTSPPPEFKIQRHEVYGVAMITINS	414
UY	461 VTAIGADSSSKRIALHNLAVLAWVIAKSGSSQGSQSMRLANLMLSHVNAS	420
DB	414 VTAIGADSSSKRIALHNLAVLAWVIAKSGSSQGSQSMRLANLMLSHVNAS	473
UY	421 EHTLNMKGVNVVYVYVIMNANVPGVYSSSTGSGSPKDESKYSGVSGVSG	477
DB	474 EHTLNMKGVNVVYVYVIMNANVPGVYSSSTGSGSPKDESKYSGVSGVSG	530
RESULT 4	ESR2_RAI	
ID	ESR2_RAI	STANDARD: PRT: 530 AA:

AC	052986: 070195: 035285: 055015: 035284: 098185:	
DI	15-JUN-1999 (Ref. 48, Created)	
DI	15-JUN-1999 (Ref. 48, last sequence update)	
DI	16-OCT-2001 (Ref. 40, last annotation update)	
DE	Estrogen receptor beta (ER-beta).	
DE	ESP2 of NP742 OF EPRBA.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
FN	[1]	
PC	SEQUENCE OF 46-530 FROM N.A. (ISOFORM BETA-1).	
PC	STRAIN Sprague-Dawley; Tissue: Prostate;	
KX	MDLINE=76234066; PubMed=8650195;	
FA	Kutiper G.G.T.M., Emarik E., Pollio Huikko M., Nilsson S.,	
FA	Guilansson J. A.;	
ET	Cloning of a novel receptor expressed in rat prostate and ovary.*	
RL	Proc. Natl. Acad. Sci. U.S.A. 93:5925-5930(1996).	
RN	[2]	
RP	SEQUENCE OF 46-530 FROM N.A. (ISOFORM BETA-2).	
PC	STRAIN Wistar; Tissue: Ovary;	
RX	MDLINE=98262942; PubMed=9600883;	
FA	Mariyama K., Esch H., Sasaki Iwaka H., Kanou H., Shimaya E.,	
RA	Hashimoto S., Kato S., Kawashima H.;	
ET	*A novel isoform of rat estrogen receptor beta with 18 amino acid	
ET	insertion in the ligand binding domain as a putative dominant	
EL	negative regular of estrogen action.*	
EL	Biochem. Biophys. Res. Commun. 241:142-147(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A. (ISOFORM BETA-2).	
PC	STRAIN Wistar; Tissue: Prostate;	
RA	Aldridge T.C.;	
ET	*Tissue specific responses to estrogen: an explanation based on	
ET	differential activation of multiple estrogen receptors with different	
ET	estrogen response elements.*	
EL	Submitted (May 1997) to the EMBL/GenBank/EBI databases.	
RN	[4]	
RP	SEQUENCE OF 46-530 FROM N.A. (ISOFORMS BETA-1; BETA-2; DELTA-3).	
PC	STRAIN Sprague-Dawley;	
RY	MDLINE=9815095; PubMed=9492041;	
FA	Petersen D.N., Tkalecovic G.T., Koza-Taylor P.H., Turf I.G.;	
RA	Brown T.A.;	
ET	*Identification of estrogen receptor beta2, a functional variant of	
ET	estrogen receptor beta expressed in normal rat tissues.*	
EL	Endocrinology 139:1082-1092(1998).	
RN	[5]	
RP	PARTIAL SEQUENCE FROM N.A. (ISOFORM BETA-1/DELTA-4).	
PC	STRAIN Sprague-Dawley; Tissue: Brain;	
FA	Prico R., Hauda R.J.;	
ET	*A novel splice variant of estrogen receptor beta found in rat	
ET	brain.*	
EL	Submitted (May 1997) to the EMBL/GenBank/EBI databases.	
CC	-1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF ER-	
CC	ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.	
CC	ISOFORMS BETA-1/DELTA-3 AND BETA-2/DELTA-4 ARE UNABLE TO BIND DNA	
CC	AND ACTIVATE TRANSCRIPTION DUE TO THE TRUNCATION OF THE DNA	
CC	BINDING DOMAIN. ISOFORM BETA-2 SHOWS LOSS OF LIGAND BINDING	
CC	AFFINITY AND SUPPRESSES ER ALPHA AND ER BETA-1 MEDIATED	
CC	TRANSCRIPTIONAL ACTIVATION AND MAY ACT AS A DOMINANT NEGATIVE	
CC	REGULATOR OF ESTROGEN ACTION.	
CC	1	
CC	SEQUENCE: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER	
CC	BETA-2.	
CC	1	
CC	ALPHA. CAN ALSO FORM HETERODIMERS BETWEEN ISOFORMS BETA-1 AND	
CC	BETA-2.	
CC	-1- SUBCELLULAR LOCATION: Nuclear.	
CC	-1- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: BETA-1 (SHOWN HERE),	
CC	BETA-1/DELTA-3, BETA-1/DELTA-4, BETA-2, AND BETA-2/DELTA-4 ARE	
CC	PRODUCED BY ALTERNATIVE SPLICING.	
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN PROSTATE, OVARY, UTERUS, LIVER,	
CC	KIDNEY, FAT, BONE, BRAIN, UTERUS AND TESTIS.	
CC	-1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,	
CC	A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID BINDING DOMAIN.	
CC	-1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.	

```

10  SSSQVH TSKKKN CHAPVKEL ISTSPQI VITL EAPPNV NMSRSMPTAS 29
20  MNNST ELAKEL VM ISMAE I DVEEI SL PQVL ESCEWVI MLIMRS IDHR7K 40

```

```

DR      HSP; p03472; 1H2O.
DR      InfctPro; PR000536; hormone_rec_14.
DR      InfctPro; PR001628; znf_6870rd.
DR      Ptm; PR00104; hormone_rec; 1.

```


the European Bioinformatics Institute. There are no restrictions on the use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

00' Actinopterygii; Neopterygii; Teleostei; Euteleostei;
01' Protacanthopterygii; Salmoniformes; Salmonidae; *Microhydrinus*.

FT DNA BIND 182 247 NUCLEAR RECEPTOR TYPE.
 FT ZN FING 182 202 C4 TYPE.
 FT ZN FING 218 242 C4 TYPE.
 FT DOMAIN 248 653 STEROID-BINDING.
 SD SEQUENCE 673 AA: 74680 MM: 2849417577E001 CR664:
 Query Match 53.78% Score 1454.53 DB 1: Length 673:
 Best Local Similarity 55.98% Prod. No. 5_50_104:
 Matches 282 Conservative 68 Mismatches 111 Indels 54 Gaps 15:
 Y 1 MNSSTISNVNLEKGR--QTTSNVIMPTGH--LSTLVYHROLSHLYAEV--KSPWC 55
 DB LVSSTP-TIN--SPSSLSLSPSAWPSNHHTMTSLTCHPESIVYNNPSPHAWL 133
 Y 56 FAR-----SLRETLVNNRETLKRRYSG-----NRCASPVYGGSKRIAHCAVCSYA 103
 DB ESAAHSTINASSSTIG--NKSIVYKSPREVEHMMSSISCAV--GKADMHCAVCHYA 189
 Y 104 SYHYVAVWSGTRKAFKRSYOGHNYGPATNQTIDKPKRSQVAFELKRYEYVAVK 164
 DB SYHYVAVWSGTRKAFKRSYOGHNYGPATNQTIDKPKRSQVAFELKRYEYVAVK 249
 Y 164 GYSRREKCYRIYKRS-----ADQIHCAKAKRSGR-----AP--RYKE 204
 DB GYSRREKCYRIYKRS-----ADQIHCAKAKRSGR-----AP--RYKE 204
 Y 250 GYVREKCYRIYKRS-----ADQIHCAKAKRSGR-----AP--RYKE 309
 DB GYVREKCYRIYKRS-----ADQIHCAKAKRSGR-----AP--RYKE 309
 Y 205 LLLATSPQVATLEAEPEYV--LSTSAFETASMMSTLAKLAKVHMLSWAKI 264
 DB LLLATSPQVATLEAEPEYV--LSTSAFETASMMSTLAKLAKVHMLSWAKI 264
 Y 410 VHSVAMSGEEFESPMDELPELYDEQKKEFTASMMSTLAKLAKVHMLSWAKI 409
 DB VHSVAMSGEEFESPMDELPELYDEQKKEFTASMMSTLAKLAKVHMLSWAKI 409
 Y 264 PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 424
 DB PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 424
 Y 470 PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 479
 DB PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 479
 Y 424 PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 492
 DB PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 492
 Y 440 PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 489
 DB PCEVESSTPGVWLEFSAWVIMMIMMSTHRSKIFATGLVLDREK--VRS--FT 489
 Y 484 ALVAVYAKSSTISGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 442
 DB ALVAVYAKSSTISGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 442
 Y 490 ALVAVYAKSSTISGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 549
 DB ALVAVYAKSSTISGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 549
 Y 443 NALVDEKSSITGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 475
 DB NALVDEKSSITGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 475
 Y 550 DANISSI--GSGSSSSSSSEYSDIOHYVQ 577
 DB DANISSI--GSGSSSSSSSEYSDIOHYVQ 577
 RESULT 16
 ESR2_1CTD0
 110 ESR2_1CTD0 STANDARD: PR1: 575 AA.
 AC CYTAKI:
 DT 16 OCT 2001 (Recl. 40, created)
 DT 16 OCT 2001 (Recl. 40, last sequence update)
 DT 16 OCT 2001 (Recl. 40, last annotation update)
 DE ESTROGEN RECEPTOR BETA (ER beta).
 GN ESR2 OR NR1A2.
 OS Tetraodon punctatus (Channel catfish).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
 OC Tetraodonti; Tetraodon.
 OX NCBI TaxID 7998;
 RN 111
 RP SEQUENCE FROM N. A.
 RC TISSUE ORGANS:
 RX MEDIANE 20218982; PubMed 10754576;
 RA Xia Z., Gao W.L., Chang X., Liaoquan D., Pei J., Maile A.G.,
 RA Desmote L.D.;
 RT Phylogenetic sequence analysis, recombinant expression, and tissue
 RT distribution of a channel catfish estrogen receptor beta.*
 RL Gen. Comp. Endocrinol. 118:139-149(2000).
 CC -1 FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 CC ER ALPHA, AND ACTIVATES EXPRESSION OF PROTEIN GENES CONTAINING
 CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.

CC -1 SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER
 CC ALPHA (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: Nucleol.
 CC -1 TISSUE SPECIFICITY: OVARY AND TESTIS.
 CC -1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC -1 NUC SUBFAMILY.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC EMBL: AF185568; AAF6157.1;
 DB HSSP: P04372; HCV.
 DB RefSeqPro: IP0000536; Hormone_Recl19.
 DB RefSeqPro: IP001723; Steroid_Receptor.
 DB RefSeqPro: IP001628; Znf_Castrolid.
 DB Pfam: PF00104; hormone_rec_1.
 DB Pfam: PF00105; Znf_C4_1.
 DB PRINTS: PR00898; STROKRENER.
 DB PROSITE: PR00047; STROIDRENER.
 DB PROSITE: PR00047; Znf_Castrolid_1.
 DB SMART: SM00430; HOLL_1.
 DB SMART: SM00430; ZNF_C4_1.
 DB PROSITE: PR00031; NUCLEAR_RECEPTOR_1.
 KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
 KW Zinc finger; Steroid binding.
 KW Modulating.
 FT DNA_BIND 161 226 NUCLEAR RECEPTOR TYPE.
 FT ZN FING 161 181 C4 TYPE.
 FT ZN FING 197 221 C4 TYPE.
 FT DOMAIN 227 575 STEROID-BINDING.
 SD SEQUENCE 575 AA: 63907 MM: 638666ABR624E CR664:
 Query Match 53.28% Score 1442.53 DB 1: Length 575:
 Best Local Similarity 55.98% Prod. No. 4_40_104:
 Matches 285 Conservative 58 Mismatches 102 Indels 65 Gaps 14:
 Y 1 MNSSTISNVNLEKGR--QTTSNVIMPTGH--LSTLVYHROLSHLYAEV--KSPWC 57
 DB LVSSTP-TIN--SPSSLSLSPSAWPSNHHTMTSLTCHPESIVYNNPSPHAWL 109
 Y 58 KSEHETLVNRETLK--KVSNCASPVGNS-----KRIAHCAV 99
 DB KSEHETLVNRETLK--KVSNCASPVGNS-----KRIAHCAV 99
 Y 110 K--PHISGKSTLSKGLAK--PEKIDKALADIDVSSSSAVVVRDIDHIVV 164
 DB K--PHISGKSTLSKGLAK--PEKIDKALADIDVSSSSAVVVRDIDHIVV 164
 Y 100 SYDASVHYGWSGTRKAFKRSYOGHNYGPATNQTIDKPKRSQVAFELKRYEYV 159
 DB SYDASVHYGWSGTRKAFKRSYOGHNYGPATNQTIDKPKRSQVAFELKRYEYV 159
 Y 165 HDVASVHYGWSGTRKAFKRSYOGHNYGPATNQTIDKPKRSQVAFELKRYEYV 224
 DB HDVASVHYGWSGTRKAFKRSYOGHNYGPATNQTIDKPKRSQVAFELKRYEYV 224
 Y 160 GMYKSGRPR--CYRIYK--GSAAGGGA--GRAPNSG 196
 DB GMYKSGRPR--CYRIYK--GSAAGGGA--GRAPNSG 196
 Y 225 GMYKSGRPR--CYRIYK--GSAAGGGA--GRAPNSG 284
 DB GMYKSGRPR--CYRIYK--GSAAGGGA--GRAPNSG 284
 Y 197 GMYKSGRPR--CYRIYK--GSAAGGGA--GRAPNSG 255
 DB GMYKSGRPR--CYRIYK--GSAAGGGA--GRAPNSG 255
 Y 285 G-----FSLSPQAVYVCLFAEPVUTLKG--MKKPPYPSVMSSTIGALRELYL 435
 DB G-----FSLSPQAVYVCLFAEPVUTLKG--MKKPPYPSVMSSTIGALRELYL 435
 Y 256 MISWAKTRGVVSLTPVGLASS--WAVVIMMIMMSTHRSKIFATGLVLDREK--VRS 415
 DB MISWAKTRGVVSLTPVGLASS--WAVVIMMIMMSTHRSKIFATGLVLDREK--VRS 415
 Y 436 MISWAKTRGVVSLTPVGLASS--WAVVIMMIMMSTHRSKIFATGLVLDREK--VRS 495
 DB MISWAKTRGVVSLTPVGLASS--WAVVIMMIMMSTHRSKIFATGLVLDREK--VRS 495
 Y 416 GVSGLLETHMLATSPREKLIOHREYV--KAMILLNNSVPIVATV--DASVSKLA 474
 DB GVSGLLETHMLATSPREKLIOHREYV--KAMILLNNSVPIVATV--DASVSKLA 474
 Y 396 GVSGLLETHMLATSPREKLIOHREYV--KAMILLNNSVPIVATV--DASVSKLA 455
 DB GVSGLLETHMLATSPREKLIOHREYV--KAMILLNNSVPIVATV--DASVSKLA 455
 Y 475 HLNATVAVVAVASGLSSGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 444
 DB HLNATVAVVAVASGLSSGSGSAPLANIMLSVPIANSKMEHLLNMR--KNVAVYDLEEM 444

456 RLLNATVIALVMAISTGISTGQASAPLHML SHIPHLNPMTHHSSMKNVYV 516
 445 YLLLEMLNAHVE: RCKSSITGSESDPA 462
 516 YLLLEMLDANMAQSHVSTGCTDPTPA 545

RESULT 17
 ERR2_CARAU STANDARD: PRT: 610 AA.
 AC 091A19;
 DT 16-OCT-2001 (rel. 40, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Estrogen receptor beta-2 (EP-beta-2)
 DE Carassius auratus (Goldfish)
 DE Elkarayot; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 DE Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 DE Cyprinidae; Carassius;
 DE Nhl_TaxID-7957;
 RN 11)
 RP SPOUNCE FROM N.A.
 RC TISSUE: Brain, and pituitary;
 RX MEDLINE-20246301; PubMed 10786629;
 RA Ma C.H., Dong K.W., Yu K.L.;
 RT cDNA cloning and expression of a novel estrogen receptor beta subtype
 in goldfish (Carassius auratus).
 RL Biochem. Biophys. Acta 1490:145-152(2000).
 CC 1 FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
 CC SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 ALPHA (BY SIMILARITY).
 CC 1-1 SUBCELLULAR LOCATION: Nucleus.
 CC 1-1 TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN PITUITARY.
 CC 1-1 TELECEPHALON AND HYPOTHALAMUS AS WELL AS IN THE LIVER.
 CC 1-1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC 1-1 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC 1-1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.

 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see: <http://www.isb.ch/permissions>
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: AF17465; AAF1770 1.
 DR HSSP: P03372; HRC;
 DR InterPro: IPR000536; Hormone_rec_119.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; Zn1_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; zn-c4_1.
 DR PRINTS: PR00348; STEROIDRMNER.
 DR PRINTS: PR00047; STEROIDFNGR.
 DR ProDom: PD000035; Zn1_C4steroid_1.
 DR SMART: SM00430; HOLL_1.
 DR SMART: SM00399; Zn1_C4_1.
 DR PROSITE: PS00041; NUCLEAR_RECEPTOR_1.
 DR RECEPTOR: Transcription regulation, DNA binding, Nuclear proteins
 KW Zinc-finger, Steroid binding
 FT DNAME 1 170 MODULATING.
 FT DNAME_BIND 171 246 NUCLEAR RECEPTOR TYPE.
 FT ZN_FING 171 191 C4 TYPE.
 FT ZN_FING 207 241 C4 TYPE.
 FT DMAIN 237 588 STEROID-BINDING.
 SU SEQUENCE 610 AA; 67851 MW; D29F2C0FC165G67 CRC64;

Query Match: 52.7%; Score 1328; DR 1; Length 610;
 Best Local Similarity: 55.5%; Pctd. No. 7.5e-102.

Matches: 277; Conserved: 61; Mismatches: 199; Gaps: 52; Gaps: 9;
 25 VLMPT-FOGL-----SPLVYHROLSHLYAFQKSPWLEAKS-LEHTLVNRETLKKV 75
 96 LPMPPHPTUSTI HCPSPILAYRPIHATHTWDAKTHINQSSSVI THA - --KLTGQGL 150
 76 SGNPVASPTGQSKDAHFCVCSIVASGTHYGVWSGCKAPFKKSTGSHNYTAT 135
 151 LQGLGILNPSGLIKGIDHFCVACHDYASGVHWSGCKAPFKKSTGSHNYTAT 210
 136 NQCTIRKPKRSQAQCH KCFYFVNVKQSSPPRCYPIVWPSASAPQI HFCGKAKS 195
 213 NQCTIRKPKRSQAQCH KCFYFVNVKQSSPPRCYPIVWPSASAPQI HFCGKAKS 269
 196 GCHAPPELLIA-----LSPQVTLLEAFPHLSRI-SAPFT 247
 270 KQISQILDELFSTHETGLHAGQNGQNSHPQVNLHAFQVLEHLEKRYT 329
 238 EASMMSTLKAKELVHMSNAKKIPEVVELSLDQVKLESCMEVLMKMLKSLDH 297
 330 EASMMSTLNIAKELVHMSNAKKIPEVVELSLDQVKLESCMEVLMKMLKSLDH 389
 298 PCKTIFAPPIVLPPEPKVPELLEPIPMI LATTSEPPPI KIQKPYIVKAMII NSSM 357
 390 PCKTIFSPQIKINPEKTVVPEIMEIPEMLLATTSEPPPIKIQKPYIVKAMII NSSM 449
 458 YPLVAT-EDASSKRLAHLLNATVIALVWIAKSGISSQOSMRILMLSHVRIAS 416
 450 GSRVQVLPDVSQKVALRLDSVIALVWIAKSGISSQOSMRILMLSHVRIAS 509
 417 NKMTEHILMMKKNVAVVILLMLNAHVIKGNSSIIQSPS----- 460
 510 NKTPEHLSMKRKNVAVVILLMLNAHVIKGNSSIIQSPS----- 460
 461 --PAFNSKSPSQNPQSO 477
 570 QQPALESNDTRHSPQAE 588

RESULT 18
 ERR2_SPAU STANDARD: PRT: 559 AA.
 ID ESR2_SPAU
 AC Q9W6M2;
 DT 16-OCT-2001 (rel. 40, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Estrogen receptor beta (EP-beta)
 DE Carassius auratus (Goldfish sea bream).
 DE Sturnus vulgaris (Common starling).
 DE Actinopterygii; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 DE Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 DE Cyprinidae; Sparus;
 DE Nhl_TaxID-8175;
 RN 11)
 RP SPOUNCE FROM N.A.
 RC TISSUE: ovary;
 RA Sotoyori S., Power D.M., Olsson P.-E., Canario A.V.M.;
 RT "Two estrogen receptors expressed in teleost fish: cDNA cloning,
 RT characterization and tissue distribution."
 RL submitted (MAR-1999) to the EMBL/Genbank/JDDB databases.
 CC 1-1 FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 CC ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
 CC 1-1 SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC 1-1 ALPHA (BY SIMILARITY).
 CC 1-1 SUBCELLULAR LOCATION: Nucleus.
 CC 1-1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC 1-1 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC 1-1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.

 CC This SWISS-Prot entry is copyright. It is produced through a collaboration

DE Fast reactor reactor gamma (EIR gamma) .
CN FBR or MRA3 .

197 JOURNAL OF THE INTERNATIONAL ASSOCIATION OF AGRICULTURAL ECONOMISTS 276


```

182 ADPQLHSGAKAKSGGHAARVREKILLDAISPELVITLLAEPPHV-LISRESAFLEAS 240
183      :      :      :      :      :      :      :      :      :
257 LITSGRTDSSSVLTGAVAVSLANPQSPALTSFQILERLEADPPREYLLMKKKLFEAK 416
258      :      :      :      :      :      :      :      :      :
241 MMSKLTFLAKRELVAHMSWAKKIPCEVEISTFQVPRITFSCMVEVIMKIMKSTIDHKK 300
242      :      :      :      :      :      :      :      :      :
317 VMSSLINLAKRELVAHMSWAKKIPCEVEISTFQVPRITFSCMVEVIMKIMKSTIDHKK 476
318      :      :      :      :      :      :      :      :      :
301 LIFAPQVLDPQPCQVEVITLFLPMLLATTSPEPELKQREYLVVPMILNLSNAPYL 260
302      :      :      :      :      :      :      :      :      :
377 LVSPEVSLSPFESVQVCFAPFPEMLLATTSPEPELKQREYLVVPMILNLSNAPYL 433
378      :      :      :      :      :      :      :      :      :
361 VITQIADSSPFIATLNAVITQAVVAVAKSTLSQVQSPMLANLMLLSHVBANFPCDM 420
362      :      :      :      :      :      :      :      :      :
434 --CLASSSESSKILRLDAVTLALVSAIGKTLVSPFQDTPLAHMLLSHLEHNSKQM 491
435      :      :      :      :      :      :      :      :      :
421 FHLINMKCKNVAVVAVITFEMI NAHVI PQCKSSITGSPSPAPHS----- 466
422      :      :      :      :      :      :      :      :      :
492 LHLHCKMKMKMNVLTLLLEMLLAHIMSHSLPHSPPEPEPELVQALATVHSSSSGTSY 551
493      :      :      :      :      :      :      :      :      :
UY 466 -----KSKENSQNP 475
UY      :      :      :      :      :      :      :      :      :
DB 552 TWTSSESGAATFQ 565

RESULT 20
ESR2_COREN1 STANDARD: PRT: 557 AA.
AC Q9YH42:
BT 16-OCT-2001 (Ref. 40, Created)
BT 16-OCT-2001 (Ref. 40, Last sequence update)
BT 16-OCT-2001 (Ref. 40, Last annotation update)
DE Estrogen receptor beta (ER-beta).
GN ESR2 OR NR4A2.
OS Opoechromis niloticus (Niloticus tilapia) (Tilapia nilotica)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Rasbora; Teleostei; Euteleostei; Neoglossini;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Clariidae; Opoechromis;
OC NR1_TaxID 8128;
RN [1]
RA SEQUENCE FROM N.A.
RA Chapo X T. Koyachi T. Togo T. Yoshitaka Y. Hiroaki I.
RA Kajima H., Nakamura M., Nagahama Y.
BT "CDNA sequence of tilapia type beta estrogen receptor ";
BT Submitted (OCT-1996) to the FMB/Jensen/K/0081 database.
CL
CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
CC ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
CC ALPHA (BY SIMILARITY).
CC -1- ALPHABETIC LOCATION: NUCLEUS.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: REFERENCES TO THE NUCLEAR HORMONE RECEPTOR FAMILY
CC NR3 SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the FMB outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See: http://www.sdb.ch/en/500).
CC or send an email to licenses@sdb.ch).
CC
DB EMBL: U75605; AAD00246.1; -
DB HSSP: P03372; IERR.
DB InterPro: IPR000536; Hormone_rec_119.
DB InterPro: IPR001234; Steroid_receptor.
DB InterPro: IPR001628; Zn_finger_4steroid.
DB Pfam: PF00104; hormone_rec_1.
DB Pfam: PF00105; zf-c4; 1.

```

```

DB PRINTS: PR00198; STERHORMNER.
DB PRINTS: PR00047; STERHORMNER.
DB PRINTS: PR00000; Zn_finger_4steroid.
DB SMART: SM00430; HOLL1.
DB SMART: SM00439; ZOF_C4; 1.
DB SMART: SM00041; NUCLEAR_RECEPTOR_1.
DB Receptor, Transcription regulation; DNA-binding; Nuclear protein;
DB Zinc-finger; Steroid-binding.
DB DOMAIN 1 154
DB DNA_BIND 155 220 NUCLEAR RECEPTOR TYPE.
DB ZN_FING 155 175 C4-TYPE.
DB ZN_FING 194 215 C4-TYPE.
DB DOMAIN 221 557 STEROID-BINDING.
DB ST000000 557 AA; 61779 MW; 82847928-196174; ORF64;

Query Match 51.6%; Score 1301; DB 1; Length 557;
Best local similarity 55.9%; Pred. No. 1; 1e-99;
Matrix: 27; Conserved: 60; Matches: 115; Indels: 46; Gaps: 12;

UY 14 GGR-----KQTSPLNVLWFLKQD-SPLVNH-KQLSHLYAFKSLWLEAKSLHTL 64
UY      :      :      :      :      :      :      :      :      :
DB 73 GGRSTSECVSHQSLASLEPSSHVGKPTLTHCPQKSGQSGASQTLW-----DSV 126
DB      :      :      :      :      :      :      :      :      :
UY 65 EVNPFETLPRVSNP-CASPVLTSPSKRIAD-CANV-SQASLYNVAWMS-TG-CAPFST 124
UY      :      :      :      :      :      :      :      :      :
DB 127 TTKSVRR---SQSPFSKVSQKADLHYCAVCAVDASVYNGVMS-PGKAPFKST 183
DB      :      :      :      :      :      :      :      :      :
UY 125 GGNHVTGVAQNC-ETTNKWPSSQVAPLPRVYV-SMVKSPSPRQVAVLVKPSADK 184
UY      :      :      :      :      :      :      :      :      :
DB 184 GGNHVTGVAQNC-ETTNKWPSSQVAPLPRVYV-SMVKSPSPRQVAVLVKPSADK 241
DB      :      :      :      :      :      :      :      :      :
UY 185 GGRACK-APRSGMNPVNH---LDMASPEQVLTTPAPRNV-LSPRQVTPFA 239
UY      :      :      :      :      :      :      :      :      :
DB 242 PLSGKTAAPKPTTPAPRNSINPKRPATTPQPTLPETAPRPTVTVTKAPRTVA 301
DB      :      :      :      :      :      :      :      :      :
UY 240 MMSKLTFLAKRELVAHMSWAKKIPCEVEISTFQVPRITFSCMVEVIMKIMKSTIDHKK 299
UY      :      :      :      :      :      :      :      :      :
DB 302 VMSSLINLAKRELVAHMSWAKKIPCEVEISTFQVPRITFSCMVEVIMKIMKSTIDHKK 361
DB      :      :      :      :      :      :      :      :      :
UY 300 KLFAPQVLDPQPCQVEVITLFLPMLLATTSPEPELKQREYLVVPMILNLSNAPYL 458
UY      :      :      :      :      :      :      :      :      :
DB 462 KLFAPQVLDPQPCQVEVITLFLPMLLATTSPEPELKQREYLVVPMILNLSNAPYL 421
DB      :      :      :      :      :      :      :      :      :
UY 459 EVNPFETLPRVSNP-CASPVLTSPSKRIAD-CANV-SQASLYNVAWMS-TG-CAPFST 419
UY      :      :      :      :      :      :      :      :      :
DB 422 SSSWSGDLDSRSKILLDAITDALVLAIGKTLVSPFQDTPLAHMLLSHLEHNSKQM 481
DB      :      :      :      :      :      :      :      :      :
UY 419 GMPHINMKCKNVAVVAVITFEMI NAHVI PQCKSSITGSPSPAPHS-----KSKENSQNP 474
UY      :      :      :      :      :      :      :      :      :
DB 482 GMPHINMKCKNVAVVAVITFEMI NAHVI PQCKSSITGSPSPAPHS-----KSKENSQNP 541
DB      :      :      :      :      :      :      :      :      :

RESULT 21
ESR1_COREN1 STANDARD: PRT: 587 AA.
AC Q91250:
BT 15-JUL-1999 (Ref. 38, Created)
BT 15-JUL-1999 (Ref. 38, Last sequence update)
BT 15-JUL-1999 (Ref. 38, Last annotation update)
DE Estrogen receptor (ER) (Estadiol receptor) (ER-alpha).
GN ESR1 OR NR3A1 OR ESR.
OS Oryzias latipes (Zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Aves; Neophalcon; Psittaci; Amniota; Psittacidae;
OC Estrildidae; Taeniopygia.
OC NR1_TaxID 59729;
RN [1]
RA SEQUENCE FROM N.A.
RA ISSUE: Hypothalamus;
RA MEDLINE: 97164508; Pubmed-9010328;
RA Jacobs F.C., Arnold A.P., Campagnoni A.T.;
RA "Zebra finch estrogen receptor cDNA cloning and mRNA expression."
RA J. Steroid Biochem. Mol. Biol. 59:135-145(1996).

```


[illegible]


```

OR PRINTS: PR00047: STEROIDRECEP.
OR PRODOM: PD000035; Znf_C4steroid; 1.
OR SMART: SM00430; HOL1; 1.
OR SMART: SM00399; ZNF_C4; 1.
OR PROSITE: PS00041: NUCLEAR RECEPTOR; 1.
OR RECEPTOR: Transcription regulation, DNA-binding, Nuclear protein.
KW Zinc-finger, Steroid-binding, Modulating.
FT DNABIND 1 179 MODULATING.
FT DNABIND 180 245 NUCLEAR RECEPTOR-TYPE.
FT ZNFING 180 200 C4-TYPE.
FT ZNFING 216 240 C4-TYPE.
FT DOMAIN 246 302 HINGE.
FT DOMAIN 303 586 STEROID-BINDING.
SU SEQUENCE 586 AA: 66080 MM: 08K077ER006P00BF CR064;

Query Mapch 47 58; Score 1199.5; OR 1; Length 586;
Best local similarity 52.08; Pred. No. 2,8e-91;
Matches 24; Conservative 76; Mismatches 117; Idols 31; Gaps 9;

UY 22 SPVWMTPTGSLSPVWVR-QLSHTYAPFYKSPWCFARSLSEHTLVNPEILKKYSCNKG 80
DB 101 SPVWFLAKLPDLSPTFHMQQVYVLESQUTFAVFAAPPTF-YRSSINRPGSGPFP 159
UY 81 ASPVTPTPT-----SKPIAFECAG-SDYASVYGVGWSGEGCAPEKSTGCHRYCTPAT 135
DB 160 MSSANIKGKPSMSNATFYCAV-SYVASYGHHGVMSDEGCAPEKSTGCHRYCTPAT 219
UY 166 NGTTTTPNPKS-QN-PIPE-YV-VWVR-VSEPTF-VYPTVFEGSAGTGG-PCVAV- 195
DB 220 NGCTIDKNNRSCAGNFKRYVYVIMKSNKSTKSTKSTKSTKSTKSTKSTKSTKSTK 274
UY 196 GCHAA-----PPVPELIDP-----ALSDQVLTLEAPRHYVLTSPSA-PPFPASM 241
DB 275 VESSTASTATWNP-VKSKSTSPVSLTAEQISAKMRAATVSTSGS-TPQSTAVM 444
UY 242 MMSITKLADKELVMMISMAKTPGVVELSPDQVLESCMWLVLMKGMMSIHTPKL 401
DB 335 MTLINLADRELHMMIMAKRYGVEDLTIHQVHLPCVAMIELMVLIMBSVHPCKL 494
UY 302 IFAPDVLDRDEKCVKSLLEIFDMLATTSPPREKLKQKELCVKAMILLNSKMYLV 361
DB 395 SFANILDLDRNCPGVVVFEDFMTVTATFEPKMP-PEEFVETKSTKSTKSTKSTKST 454
UY 362 TATQDAUSKRLAH-LINAVTALVWVIAKSGSSGOSMRANLMLLSHVHAKNKM 420
DB 455 SSTLESTPTPTLHTLHPKTPPTVHPMAKSGSLGQGPPLAQTLTTHSHIPMSNPM 511
UY 421 EHLNKKCKNVVYVDLLEMLNAVLKCKKSTTSGSPADHSK 467
DB 515 EHYMKCKNVVPLDILLEMLAHPTTKTKTKTGF----PDSGS 557

RESULT 27
ESRL_MOUSE STANDARD: PKT; 599 AA.
ID ESRL_MOUSE STANDB: PKT; 599 AA.
AC P19785; Q9QY51; Q9QY52;
DE 01-FEB-1991 (Rel. 17, Created)
DE 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Estrogen receptor (ER) (Estrogen receptor) (ER-alpha).
DE ERK1 OR NR3A1 OR ESR OR ESRB OR ESRP
DE Mus musculus (Mouse).
DE Eukaryote; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
DE NCBI_Taxid=10090;
OR SEQUENCE FROM N.A.
OR TISSUE: uterus;
OR MEDLINE=91042558; PubMed=2484714;
OR White R.; Lewis J.A.; Needham M.; Han T.; Parker M.;
OR "Structural organization and expression of the mouse estrogen
OR receptor";
OR Mol. Endocrinol. 1:735-744(1987).

```

```

KN 121
KN SEQUENCE OF 269-599 FROM N.A.
KN STAIN-STRUT; and R10 S/1; TISSUE: Spleen;
KN Ma R.Z.; Teuscher G.;
KN "Screening for candidate genes of mouse autoimmune diseases";
KN selected (Jan 1999) to the FMI/Genbank/DBM databases.
KN [3]
KN CARBOHYDRATE-LINKAGE SITE SER-575.
KN MEDLINE=97153020; PubMed=899954;
KN Jiang M.S.; Hart G.W.;
KN "A subpopulation of estrogen receptors are modified by O-linked
KN N-acetylglucosamine";
KN J. Biol. Chem. 272:2421-2428(1997).
KN [4]
KN CARBOHYDRATE-LINKAGE SITES SER-10; THR-50 AND SER-575.
KN MEDLINE=212447; PubMed=11256831;
KN Cheng X.; Hart G.W.;
KN "Oligosylation of the murine estrogen receptor-alpha";
KN J. Steroid Biochem. Mol. Biol. 75:147-158(2000).
KN 1. FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
KN THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
KN PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
KN 2. SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
KN BETA (BY SIMILARITY).
KN 3. SUBCELLULAR LOCATION: Nuclear.
KN 4. DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
KN A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
KN 5. FURTHER INFORMATION: CYCLIN A/CYC2 (BY SIMILARITY).
KN 6. MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
KN ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
KN HORMONE-BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
KN HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCREET DNA
KN SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
KN 7. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
KN 8. SUBFAMILY.
KN -----
KN THIS SWISS-PROT entry is copyright. It is produced through a collaboration
KN between the Swiss Institute of Bioinformatics and the EMBL outstation -
KN the European Bioinformatics Institute. There are no restrictions on its
KN use by non-profit institutions as long as its content is in no way
KN modified and this statement is not removed. Usage by and for commercial
KN entities requires a license agreement (see http://www.isb-sib.ch/aboutus/
KN or send an email to license@isb-sib.ch).
KN -----
KN EMBL: M38651; AAA37580.1; --
KN EMBL: AF128231; AAF22662.1; --
KN EMBL: AF128230; AAF22661.1; --
KN PIR: A40061; A40061.
KN HSSP: P03372; HSCP.
KN TRANSFAC: T00259; --
KN GLYCOSULEDB: P19785; --
KN MDJ: M31332467; ESRL.
KN InterPro: IPK000536; Hormone_Recc_14.
KN InterPro: IPK001292; oestrgn_receptor.
KN InterPro: IPK001724; Sterhm_123.
KN InterPro: IPK001628; Znf_C4steroid.
KN Pfam: PF00104; hormone_recc_1.
KN Pfam: PF00105; zf-c4; 1.
KN Pfam: PF02159; oest_reccp_1.
KN PRINTS: PR000478; STEROIDRECEP.
KN PRINTS: PR000478; STEROIDRECEP.
KN PRODOM: PD000035; Znf_C4steroid; 1.
KN SMART: SM00430; HOL1; 1.
KN SMART: SM00399; ZNF_C4; 1.
KN PROSITE: PS00041: NUCLEAR RECEPTOR; 1.
KN RECEPTOR: Transcription regulation; DNA-binding; Nuclear protein;
KN Zinc-finger; Steroid-binding; Phosphorylation; Glycoprotein.
KN KW DNABIND 1 188 MODULATING.
KN DNABIND 189 254 NUCLEAR RECEPTOR-TYPE.
KN ZNFING 189 209 C4-TYPE.
KN ZNFING 225 249 C4-TYPE.
KN DOMAIN 255 414 HINGE.
KN DOMAIN 415 599 STEROID-BINDING.

```


Query Match: 47.4%, Score 1195, DR 1, Length 579,
 Best Local Similarity: 50.1%, Pred. No. 6,50-91;
 Matches: 252; Conservative: 74; Mismatches: 120; Indels: 42; Gaps: 14;

4 YSTD-----SNVFNEDGPGDPTSPNVLWPTPHISPIVHPQI SHI YATPCKSP 53
 44 YSNAPIDAHDPSPNSISLQSLSP-----NSPIVAVSSPSISPTM--GQANHHYFTTSTF 96
 54 WFKASLEHLLVNNRTI...KKVSNAPASVPTSPSPKAPHPVAVSVASVYH 107
 97 YVAVPSNQS--VSPETAPVTSI SPFSV--VPSN--ACACSEPMKEMPEVAVSVASVYH 151
 108 YVAVSSGCKAPKPKPSLQGHNIYTPATVCTIIONPKKSGACPIKGVYVAVYQVQSP 167
 155 YVAVSSVPRKAPKPKPSLQGHNIYTPATVCTIIONPKKSGACPIKGVYVAVYQVQSP 214
 158 REFQCYLVLPGR-----SAPDQIPKACPKSSQAHVAVVPI 206
 215 KDR--GVVLPKPKPTSTSPKASKGLEHTAPAPQDPKPTSSAVVQVCKSVI----- 268
 207 LMLSLBQVLTLEAFTHVLSRP--SAFTESMMMSILKALAKELVHMSAANL 264
 269 --SMPVGVITIPKAPPP--MPSCKVMPPTVEVIMLLISMAKELVHMAKKEP 325
 265 GFVELSLFGVRLIFSCMMVLMMLMBSIHPKELFAPRLVLPKCEVEITIEF 324
 326 GFIQISIHQVGLIFSSNMFVIMTLIMPSIHPKELFAPRLVLPKCEVEITIEF 385
 325 DMLATTSRELEKIQHKRYIVKAMITINSWYPIVATQD--AVSSPKAHINAVVPA 384
 386 DMLATVAPRPMIKLPPEFVTKATILLNSAFSPCTGTMPRLHNSAAVQMLPTIDA 445
 384 LVAVLAKSSISSQOQMLANLMLLSHVHANSKCMELLMMKKKNVPPYDILLPMIN 444
 446 LHHHNSGCSAQQSPQACIILITSHIPKSNQPMFHYSPKPVNVPVYITIPMD 505
 444 AH VLKCKSSITSGES-----PAEDSKSKET 470
 506 AHVHPDPAPFVMSQADPELFTSSSSSSS 548

RESULT 29
 ESR1_FAI
 ID ESR1_FAI STANDARD: PRT: 600 AA.
 AC P06211;
 DI 01-JAN-1988 (Ref. 06, created)
 DI 01-JAN-1988 (Ref. 06, last sequence update)
 DI 16-OCT-2001 (Ref. 40, last annotation update)
 DI Estrogen receptor (ER) (Estrogen receptor) (ER-alpha).
 GN ESR1 OR NR4A1 OR ESR OR ESR1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NX MBLTaxid: 10116;
 RN 111
 RE SEQUENCE FROM N.A.
 RA STRAIN Mistral;
 RA Muramatsu M.;
 RI 3, limited (MAR 1987) to the EMBL/Genbank/Trna databases.
 RN 121
 RE SEQUENCE FROM N.A.
 RA MEDLINE:87174780; PubMed:4031601;
 RA Koike S., Sakai M.;
 RI Molecular cloning and characterization of rat estrogen receptor
 RI cDNA*;
 RI Nucleic Acids Res. 15:2499-2513(1987).
 RN 141
 RE SEQUENCE FROM N.A.
 RA STRAIN Sprague-Dawley; TISSUE: uterus;
 RA Maggi A.M.A.;
 RI Submitted (JUN-1991) to the EMBL/Genbank/Trna databases
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR

CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC 1 SUBUNIT BINDS DNA AS A HOMODIMER, CAN FORM A HETERODIMER WITH ER-
 CC BETA (BY SIMILARITY).
 CC 1-1 SUBCELLULAR LOCATION: Nuclear.
 CC 1-1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC 1-1 EFM: FIBERCORRELATED BY CYCLIN A/CDK2 (BY SIMILARITY).
 CC 1-1 ARE THOUGHT TO BE WEAVELY ASSOCIATED WITH NUCLEAR HORMONE RECEPTORS
 CC HORMONE-BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
 CC HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA
 CC SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
 CC 1-1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NBS SUPERFAMILY.
 CC 1-1 This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC 08 EMBL: Y00102; CAA68287.1;
 CC 08 EMBL: X41096; CAA43411.1;
 CC 08 PIR: S07379; S07379.
 CC 08 HESP: P03372; HICP.
 CC 08 TRANSFAC: T00258;
 CC 08 InterPro: IPR000536; hormone_rec_119
 CC 08 InterPro: IPR001292; oestrgn_receptor.
 CC 08 InterPro: IPR001723; steroid_receptor.
 CC 08 InterPro: IPR001628; znf_c4steroid.
 CC 08 Pfam: PF00104; hormone_rec_1.
 CC 08 Pfam: PF00105; zf-c4; 1.
 CC 08 Pfam: PF02159; oest_recep_1.
 CC 08 PRINTS: PR00398; STDRHORMONER.
 CC 08 PRINTS: PR00047; STDRHORMONER.
 CC 08 Protein: P000035; Znf_C4steroid; 1.
 CC 08 SMART: SM00394; Znf_C4; 1.
 CC 08 PROSITE: PS00041; NUC_FAB_RECEPTOR; 1.
 CC 08 Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc finger, Steroid binding, Phosphorylation.
 CC FT DOMAIN 1 189
 CC FT DNA_BIND 190 255
 CC FT ZN_FING 190 210
 CC FT ZN_FING 226 250
 CC FT DOMAIN 256 315
 CC FT DOMAIN 316 600
 CC FT DOMAIN 64 71
 CC FT DOMAIN 171 174
 CC FT MOD_RES 109 109
 CC FT MOD_RES 111 111
 CC FT MOD_RES 124 124
 CC FT MOD_RES 172 172
 CC FT MOD_RES 542 542
 CC FT CONFLICT 488 488
 CC FT SEQUENCE 600 AA; 67030 MW; C9V7D8CACEV57DR CIG64;
 CC 08 15 GPERG-----TSPNVLWPTPHISPIV--VHPQI SHI YATPCKSP 53
 47.4%; Score 1194.5; DR 1; Length 600;
 Best Local Similarity: 49.0%; Pred. No. 7,50-91;
 Matches: 244; Conservative: 90; Mismatches: 109; Indels: 55; Gaps: 11;

44 YSNAPIDAHDPSPNSISLQSLSP-----NSPIVAVSSPSISPTM--GQANHHYFTTSTF 96
 54 WFKASLEHLLVNNRTI...KKVSNAPASVPTSPSPKAPHPVAVSVASVYH 107
 97 YVAVPSNQS--VSPETAPVTSI SPFSV--VPSN--ACACSEPMKEMPEVAVSVASVYH 151
 108 YVAVSSGCKAPKPKPSLQGHNIYTPATVCTIIONPKKSGACPIKGVYVAVYQVQSP 167
 155 YVAVSSVPRKAPKPKPSLQGHNIYTPATVCTIIONPKKSGACPIKGVYVAVYQVQSP 214
 158 REFQCYLVLPGR-----SAPDQIPKACPKSSQAHVAVVPI 206
 215 KDR--GVVLPKPKPTSTSPKASKGLEHTAPAPQDPKPTSSAVVQVCKSVI----- 268
 207 LMLSLBQVLTLEAFTHVLSRP--SAFTESMMMSILKALAKELVHMSAANL 264
 269 --SMPVGVITIPKAPPP--MPSCKVMPPTVEVIMLLISMAKELVHMAKKEP 325
 265 GFVELSLFGVRLIFSCMMVLMMLMBSIHPKELFAPRLVLPKCEVEITIEF 324
 326 GFIQISIHQVGLIFSSNMFVIMTLIMPSIHPKELFAPRLVLPKCEVEITIEF 385
 325 DMLATTSRELEKIQHKRYIVKAMITINSWYPIVATQD--AVSSPKAHINAVVPA 384
 386 DMLATVAPRPMIKLPPEFVTKATILLNSAFSPCTGTMPRLHNSAAVQMLPTIDA 445
 384 LVAVLAKSSISSQOQMLANLMLLSHVHANSKCMELLMMKKKNVPPYDILLPMIN 444
 446 LHHHNSGCSAQQSPQACIILITSHIPKSNQPMFHYSPKPVNVPVYITIPMD 505
 444 AH VLKCKSSITSGES-----PAEDSKSKET 470
 506 AHVHPDPAPFVMSQADPELFTSSSSSSS 548

or send an email to ncbiinfo@ncbi.nlm.nih.gov.

001 EMBL: AF255605; A012454.1; 5'
 002 EMBL: AF061275; A076954.1; 5'
 003 EMBL: AF255606; A012454.1; 5'
 004 HSSB: J04072; JEPF.
 005 InterPro: IP0000546; Hormone_rec_14.
 006 InterPro: IP0001292; oest_recep_14.
 007 InterPro: IP0001724; Stdhorm_receptor.
 008 InterPro: IP0001628; Znf_C4steroid.
 009 Pfam: PF00104; Hormone_rec_1.
 010 Pfam: PF00105; Zf_C4_1.
 011 Pfam: PF02159; oest_recep_1.
 012 PRINTS: PR00498; STROHORMNER.
 013 PRINTS: PR00477; STROHPTIMER.
 014 ProDom: PD000435; Znf_C4steroid_1.
 015 SMART: SM00400; H0L1; 1.
 016 SMART: SM00499; Znf_C4_1.
 017 Pfam: PF00041; NUCLEAR_RECEPTOR_1.
 018 Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 019 Zinc finger; Steroid binding; Alternative splicing.
 020 DOMAIN 1 186 MODULATING.
 021 DNA_BIND 187 252 NUCLEAR RECEPTOR TYPE.
 022 ZN_FING 187 207 C4 TYPE.
 023 ZN_FING 223 247 C4 TYPE.
 024 DOMAIN 253 315 HINGE.
 025 DOMAIN 316 617 STEROID-BINDING (POTENTIAL).
 026 DOMAIN 421 425 POLY-LEU.
 027 DOMAIN 512 516 POLY-THR.
 028 VASBP 577 591 POLY-THR.
 029 VASBP 1 46 MISSING (IN ISOFORM 2).
 030 VASBP 1 425 MISSING (IN ISOFORM 3).
 031 SEQUENCE 617 AA; 67670 MW; 7b208558e5ad403 CRC64;

Query Match 45.6%; Score 1151; DB 1; Length 617;
 Best Local Similarity 51.0%; Proc. No. 40-87;
 Matches 251; Conservative 75; Mismatches 110; Indels 58; Gaps 14;

001 4 YSI --PSNVTNLEKHPGRTSPNVLMPVGH.SPLVNH -----POLSHYAP 49
 002 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 003 76 YSVAPEDQEDIDPDM: SSSPVEVSSPSPELGHDAQGHAGVPPYLEPS 131
 004 50 QSNWFEAS.LHHLVYVRELAKRY--SNRFAVSTPR--GSKPAHFAVVS 100
 005 132 GTSTYKSSVLAASASKPLASVAGVAVYAVASPSGASGATGIVKELRYVSVS 191
 006 101 DYASGVHYGVWSEFKKAFKKITQSHINYICFATWCTDKNRERSVACRLKRYEVS 160
 007 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 008 192 DYASGVHYGVWSEFKKAFKKITQSHINYICFATWCTDKNRERSVACRLKRYEVS 251
 009 161 MKVGSRRKGVGLVVRGVSADPOLACAKAKSKCHAVR-----VREL----- 206
 010 252 MKKQEPKRPGRPELHNPDR--SKLKEHKGYSKAGVSGSVETLQVDSSSA 304
 011 207 LALSPQGLVTLLEAPGPHLSKPE-SAPFEASMMSLIKLAKRVLVHI 257
 012 405 LGGVAVLVVMSPPGVILLKRPPT-TLSPEKHSPPVSPETLSLTNNALGELVHI 464
 013 278 SWARKLGGVHLSTFQVRLLESWMVLMMLGMMWSLDHCKLIPAPLVLDROCKV 317
 014 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 015 364 AMARKVGFQSLSHDVOLESSWELIMIGLWISVTPKRLIPAOULLKRSRGEV 423
 016 418 ELLELEEMALITSLHPLKIQKRPVYVYVAMILNNSMYLVIA 156VASSGELAH 479
 017 424 KMAHFDRLAIVAVRRTIKKSEEPVAKAILLNSVASSSSSVETLQVDSSSA 483
 018 477 LNAVTAIVVYAKSCTSSQGSMLANILMLSHVHANSKQMEHLMMKTKNVAVYD 496
 019 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 020 484 MNTIDALVYTSQSLISVLOSRRQGLLLLSHIRHMSYKQMEHMMKKNVAVYD 543
 021 417 LLEPLNANVIR 448
 022 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 023 544 LLEPLNANVIR 555

RESULT 48
 ESR1 CONCY STANDARD: PRT: 622 AA.
 ID ESR1_NORM G0F529, G0F529, G0YGR0;
 AC P10058, G0F529, G0F529, G0YGR0;
 DT 01-Apr-1990 (Ref. 14, Unreviewed)
 DE 16-Oct-2001 (Ref. 40, Last sequence update)
 DE 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Estrogen receptor (ER) (Estrogen receptor) (ER-alpha)
 DE ESR1 OR NR4A1 OR ERK.
 GN Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OS Oncorhynchus mykiss (Chordata; Craniata; Vertebrata; Euteleostomi);
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protactinopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCIT: T010202;
 RX MEDLINE 9106824; PubMed 2210041;
 RA Pakdel F., Le Gac F., Le Gall P., Valotiato Y.;
 RT "Full-length sequence and in vitro expression of rainbow trout
 estrogen receptor cDNA."
 RL Mol. Cell. Endocrinol. 71:195-204(1990).
 RN 121
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RX MEDLINE 9106824; PubMed 2210041;
 RA Pakdel F., Le Gac F., Le Gall P., Valotiato Y.;
 RT "Full-length sequence and in vitro expression of rainbow trout
 estrogen receptor cDNA."
 RL Mol. Cell. Endocrinol. 71:195-204(1990).
 RN 121
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RX MEDLINE 20114460; PubMed 10650938;
 RA Pakdel F., Floutier R., Floutier G., Valotiato Y.;
 RT "Two estrogen receptor (ER) isoforms with different estrogen
 dependencies are generated from the trout ER gene."
 RL Endocrinology 141:571-580(2000).
 RN 141
 RP SEQUENCE OF 150-574 FROM N.A. (SHORT ISOFORM).
 RX MEDLINE 9106824; PubMed 2210041;
 RA Pakdel F., Le Gac F., Le Gall P., Valotiato Y.;
 RT "Full-length sequence and in vitro expression of rainbow trout
 estrogen receptor cDNA."
 RL Mol. Cell. Endocrinol. 71:195-204(1990).
 RN 141
 RP SEQUENCE OF 279-642 FROM N.A.
 RX MEDLINE 20159854; PubMed 10696781;
 RA Matthews J.H., Zacharewski T.R.;
 RT "Differential binding affinities of ERs, hsp-70s and Aroclors with
 recombinant human, rainbow trout (Oncorhynchus mykiss) and green
 anole (Anolis carolinensis) estrogen receptors using a semi high
 throughput competitive binding assay."
 RL Toxicol. Sci. 53:426-434(2000).
 CC 1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC 2- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 BETA (BY SIMILARITY).
 CC 3- SUBCELLULAR LOCATION: Nucleus.
 CC 4- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM/ER(1) (SHOWN HERE)
 AND A SHORT FORM/ER(2); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC 5- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC 6- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
 ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
 HORMONE BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
 HORMONE-RECEPTOR COMPLEX ATTENDS TO RECRUITED DISPERSE DNA
 SEQUENCES OF STRIPED OR TRANSCRIPTIONAL START SITES.
 CC 7- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC 8- NRS SUBFAMILY.

This SWISS-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcements>)

17 GGGTTSSTNLTMPPTGHSPLVYHKOISH --- LTAPEOKSPWCEALSTHTDTPVNET 70
 17 GSSPTTPTVYSSSSGYSQSPD --- DPGDGHEDPQSSYTFEISTTPTVYSSSVNGLSASPE 74
 71 LKKKSGNNGCAS PVYHGGSKR --- KLAHPVAVSYVASGYHGVWSSGEGCKAKPKRSI 124
 75 KLTLLTSNGQSSVAAGSCVAVPE EMMANLIRYCAVSYGFASQY HGVWSSGEGCKAKPKRSI 134
 125 GYHINYPVATNGCTLDKNSPKSSVQA KLEKRYEYGMKRYNSKREKRYKLYVRQSGADE 184
 134 GYHINYPVATNGCTLDKNSPKSSVQA KLEKRYEYGMKRYNSKREKRYKLYVRQSGADE 187
 185 GHHGK ACCKAKR --- SCGH --- ARPVHETLDDMSPEOL 215
 186 KRYCTVADREKRYVGLLEHRTAPVYQGGSSSSSSSSGSGGWSGCRK --- TMPEQV 241
 216 VETLEAEVPHVILSRD --- SAPEEASMMMSITKLAOKELVIMI SWARKHGVPEVSLPD 274
 242 LTLTGAEVPE ALCSKQKVAPIYETVETVMTLT --- MAOKELVIMIAKAKRYGPEVSLPD 309
 274 QVADLESYVMEVIMKMLMKRSIDHGGKLEFADVYLDDEKRYVGLLEFEDMLATTSK 334
 401 QVADLESYVMEVIMKMLMKRSIDHGGKLEFADVYLDDEKRYVGLLEFEDMLATTSK 409
 334 FREKLOKREYLVKAMILLNSMYPIVY -ATQVADSSSKLAHLINAVTVALVWVIAKSG 392
 361 FREKLOKREYLVKAMILLNSMYPIVY -ATQVADSSSKLAHLINAVTVALVWVIAKSG 420
 494 ISSGQSSMKLANLMLLSVYVIAKSKKREHILNKKRYVAVVYVGLLEFEDMLATTSK 452
 421 ASVQDQPRQVQGLLTLSSIKMSKMPHLYSTK KKKVLYVGLLEFEDMLATTSK 476
 453 SLEGGSPVSPKSKSKSSQNPQS 476
 477 --- SSGKVAQNGHETPEST 494

RESULT 40
 ESTBL_CNFIRM
 ID ESTBL_CNFIRM STANDARD PRT 407 AA.
 AC GY1424:
 DT 16 OCT 2001 (Ref. 40, Greated)
 DT 16 OCT 2001 (Ref. 40, last annotation update)
 DE Estrogen receptor (ER) (Estrogen receptor) (ER-alpha) (Fragment)
 GS ESTBL OR NR4A1 OR ESR.
 OS Chemodiphorus uniparous (Whiplash lizard).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Sceloporiformes; Teiioidea;
 OC Teiidae; Chemodiphorus.
 OX NR1 Ptxid 37197;
 RN 111
 RP SOURCE FROM N.A.
 RC TISSUE: ovary, and kidney;
 RC MEDLINE 96103562; PubMed 7495707;
 RA Young L.L., Gaskin J., Giamber M., Galt M., Grows D.;
 RT "Ovarian sex steroid receptors: amplification, sequence and
 KJ expression analysis."
 RL J. Steroid Biochem. Mol. Biol. 55:261-269(1995).
 CC 1 FUNCTION: THE STEROID HORMONE AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC 1 SUBUNIT: BINDS DNA AS A HOMODIMER, CAN FORM A HETERODIMER WITH ER
 CC BETA (BY SIMILARITY).
 CC 1 SUBCELLULAR LOCATION: Nucleus.
 CC 1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA BINDING DOMAIN AND A C-TERMINAL STEROID BINDING DOMAIN.
 CC 1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC N/A SIMILARITY.

This SWISS-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its
 use by non profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (<http://www.ebi.ac.uk/infocentre/copyright.htm>
 or send an email to license@ebi.ac.uk).

EMBL: S79924; AAB45749.1;
 D8 HSSP: P04372; HRP.
 D8 IntOPro: IPRO00546; hormone_rec_119.
 D8 IntOPro: IPRO01723; Steroid_receptor.
 D8 IntOPro: IPRO01628; ZnF_C4steroid.
 D8 Pfam: PF00104; hormone_rec_1.
 D8 Pfam: PF00105; zt_c4_1.
 D8 PRINTS: PR00398; STEROHORMONER.
 D8 PRINTS: PR00047; STEROIDREINER.
 D8 PRODOM: PRO00055; ZnF_C4steroid_1.
 D8 SMART: SM00430; ZnF_C4_1.
 D8 SMART: SM00439; ZnF_C4_1.
 D8 PROSITE: PS00041; NUCLEAR_RECEPTOR; PARTIAL.
 D8 Recyptor: Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding.
 FT NON_TER 1
 FT DNA_BIND <1 36 NUCLEAR RECEPTOR TYPE.
 FT ZN_FING 7 31 C4-TYPE.
 FT DOMAIN 47 95 HINCE.
 FT DOMAIN 96 107 STEROID-BINDING.
 SO SPOUNCE 407 AA; MW: 10801426; pI: 6.64;

Query Match 34.8%; Score 877; DB 1; Length 407;
 best local similarity 54.8%; Prod. No. 4, 36-65;
 Matches 172; Conservative 62; Mismatches 94; Indels 26; Gaps 7;

126 GYHINYPVATNGCTLDKNSPKSSVQA KLEKRYEYGMKRYNSKREKRYKLYVRQSGADE 185
 1 GYHINYPVATNGCTLDKNSPKSSVQA KLEKRYEYGMKRYNSKREKRYKLYVRQSGADE 60
 186 LKACAKRSQ --- LKAPV --- RELLE --- ALSPQVLTLLFAHPVLT 229
 61 --- DNEHGAIVERSKSPDMPSPMLTINKNSVIALTAQVIALFAHPV VV 114
 230 SR--PSAPEEASMMMSITKLAOKELVIMI SWARKHGVPEVSLPDVRLLESYVMEVIM 287
 114 SEYDPSKPSPEASMMITLTNLAKRELYMIMAKKRYGPEVSLPDVHLLERVAMELTM 174
 288 MLLMKRSIDHGGKLEFADVYLDDEKRYVGLLEFEDMLATTSKRELAOKREYLV 447
 174 TGVWSSVPHGKLEFADVYLDDEKRYVGLLEFEDMLATTSKRELAOKREYLV 234
 348 KAMILLNSMYPIVAT QDAISSKLAHLINAVTVALVWVIAKSGISSGQSSMKLANL 406
 234 KSLILLNSGYVPLSSITKSKREKHIVLEKLTDTLLHMAKSGSLDQOHKRLAOL 294
 407 MLLSHVRIANSKGM 420
 294 LTLSSIRMSKGM 407

Search completed: June 16, 2003, 14:14:33
 Job time: 14.0675 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 16, 2003, 11:34:03 ; Search time 27.487 seconds

(without alignments)
3575.671 Million cell updates/seq

Title: US-08-R26-361a-5

Perfect score: 2532

POSTALFOSKSPSPSPSPSPSP 477

Sequence: 1 MVSSTPSNTNLESPSPOT

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 120 comparisons

Database :

1: SPITRMBL_21:
2: SP_archaea:
3: SP_bacteria:
4: SP_fungi:
5: SP_human:
6: SP_invertebrate:
7: SP_mammal:
8: SP_mhc:
9: SP_oranella:
10: SP_plant:
11: SP_rodent:
12: SP_virus:
13: SP_vertebrate:
14: SP_unclassified:
15: SP_rivius:
16: SP_bacteriap:
17: SP_archaeap:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2241	88.5	593	11	Q91286	Q91286 mus muscula
2	2141	84.9	499	6	Q95M60	Q95M60 macaca arcu
3	2058	81.6	486	6	Q95M69	Q95M69 gallitrix
4	1558.5	61.8	542	13	Q90W56	Q90W56 squilla sca
5	1469	54.3	553	13	Q90W58	Q90W58 brachydanto
6	1360	53.9	553	13	Q90W57	Q90W57 brachydanto
7	1354	53.7	552	13	Q90W59	Q90W59 brachydanto
8	1354	53.7	552	13	Q90W59	Q90W59 brachydanto
9	1295.5	51.4	562	13	Q90W75	Q90W75 brachydanto
10	1243	49.3	550	6	Q95JCO	Q95JCO canis fami
11	1211	48.0	587	13	Q90W40	Q90W40 canis fami
12	1210	48.0	581	13	Q90W49	Q90W49 canis fami
13	1181.5	46.8	569	13	Q90W59	Q90W59 brachydanto
14	1170.5	46.4	564	13	Q90W51	Q90W51 brachydanto
15	1165.5	46.2	431	6	Q95L13	Q95L13 canis fami
16	1156	45.8	578	13	Q90W40	Q90W40 canis fami

17	1141.5	45.3	620	13	Q90W56	Q90W56 brachydanto
18	1113	44.1	554	13	Q90W58	Q90W58 brachydanto
19	1099	43.6	458	13	Q90W57	Q90W57 brachydanto
20	1061.5	42.1	391	13	Q90W58	Q90W58 brachydanto
21	942.5	37.4	335	13	Q90W54	Q90W54 brachydanto
22	679	26.9	274	13	Q91984	Q91984 brachydanto
23	594.5	23.6	200	13	Q90W59	Q90W59 brachydanto
24	525	20.8	101	13	Q923X1	Q923X1 mesocriceta
25	522	20.7	484	5	Q90W59	Q90W59 drosophila
26	522	20.7	435	5	Q90W59	Q90W59 drosophila
27	517.5	20.5	159	6	Q95JCO	Q95JCO canis fami
28	504	20.1	249	6	Q95JCO	Q95JCO canis fami
29	504	20.0	710	13	Q91985	Q91985 canis fami
30	479	19.0	593	13	Q90W54	Q90W54 canis fami
31	475.5	18.9	323	13	Q90W54	Q90W54 canis fami
32	470.5	18.7	703	13	Q90W54	Q90W54 canis fami
33	461	18.3	438	13	Q90W57	Q90W57 brachydanto
34	460.5	18.3	689	13	Q90W58	Q90W58 brachydanto
35	459.5	18.2	712	13	Q90W59	Q90W59 brachydanto
36	458.5	18.2	525	6	Q95L13	Q95L13 canis fami
37	458	18.2	934	6	Q95L13	Q95L13 canis fami
38	452.5	17.9	414	11	Q90W56	Q90W56 mus muscula
39	449.5	17.8	400	5	Q90W56	Q90W56 mus muscula
40	448.5	17.8	436	5	Q90W56	Q90W56 mus muscula
41	448.5	17.8	688	13	Q90W58	Q90W58 brachydanto
42	448	17.8	790	13	Q90W48	Q90W48 xenopus lae
43	442	17.5	87	6	Q90W56	Q90W56 mus muscula
44	442	17.5	157	13	Q90W56	Q90W56 mus muscula
45	442	17.5	848	13	Q90W59	Q90W59 brachydanto
46	441.5	17.5	401	13	Q90W56	Q90W56 mus muscula
47	440.5	17.5	360	13	Q90W56	Q90W56 mus muscula
48	437	17.3	982	6	Q90W58	Q90W58 brachydanto
49	436	17.3	414	5	Q90W56	Q90W56 mus muscula
50	435	17.2	363	5	Q90W56	Q90W56 mus muscula
51	434	17.2	906	4	Q90W56	Q90W56 mus muscula
52	433	17.2	82	6	Q90W56	Q90W56 mus muscula
53	432	17.1	412	13	Q90W56	Q90W56 mus muscula
54	432	17.1	730	4	Q90W56	Q90W56 mus muscula
55	431.5	17.1	793	11	Q90W56	Q90W56 mus muscula
56	431.5	17.1	794	11	Q90W56	Q90W56 mus muscula
57	430.5	17.1	895	6	Q90W56	Q90W56 mus muscula
58	430.5	17.1	896	6	Q90W56	Q90W56 mus muscula
59	430	17.0	853	13	Q90W56	Q90W56 mus muscula
60	429.5	17.0	427	5	Q90W56	Q90W56 mus muscula
61	427.5	17.0	692	13	Q90W56	Q90W56 mus muscula
62	426	16.9	763	13	Q90W56	Q90W56 mus muscula
63	425	16.9	340	5	Q90W56	Q90W56 mus muscula
64	425	16.9	854	13	Q90W56	Q90W56 mus muscula
65	423.5	16.8	334	4	Q90W56	Q90W56 mus muscula
66	423	16.8	438	13	Q90W56	Q90W56 mus muscula
67	421.5	16.7	769	13	Q90W56	Q90W56 mus muscula
68	421	16.7	849	13	Q90W56	Q90W56 mus muscula
69	419	16.6	379	13	Q90W56	Q90W56 mus muscula
70	417	16.5	452	13	Q90W56	Q90W56 mus muscula
71	417	16.5	563	13	Q90W56	Q90W56 mus muscula
72	416.5	16.5	346	4	Q90W56	Q90W56 mus muscula
73	416.5	16.5	475	5	Q90W56	Q90W56 mus muscula
74	415.5	16.5	448	5	Q90W56	Q90W56 mus muscula
75	414.5	16.4	427	13	Q90W56	Q90W56 mus muscula
76	411.5	16.3	303	6	Q90W56	Q90W56 mus muscula
77	411	16.3	422	13	Q90W56	Q90W56 mus muscula
78	409	16.2	389	5	Q90W56	Q90W56 mus muscula
79	408	16.2	380	5	Q90W56	Q90W56 mus muscula
80	405	16.1	797	13	Q90W56	Q90W56 mus muscula
81	404.5	16.0	344	13	Q90W56	Q90W56 mus muscula
82	403	15.9	438	5	Q90W56	Q90W56 mus muscula
83	403	15.8	763	13	Q90W56	Q90W56 mus muscula
84	399	15.8	322	11	Q90W56	Q90W56 mus muscula
85	396	15.7	539	11	Q90W56	Q90W56 mus muscula
86	396	15.7	540	11	Q90W56	Q90W56 mus muscula
87	392	15.5	600	13	Q90W56	Q90W56 mus muscula
88	387.5	15.4	467	5	Q90W56	Q90W56 mus muscula
89	385	15.3	541	4	Q90W56	Q90W56 mus muscula

Matches 404: Conservative 5: Mismatches 15: Indels 0: Gaps 0

QY 1 MNSYISNVNIEGSGPOTTSNVIWPTGSHSPLVVRGQSHIYAPPOKSPWFAKSI 60
|||||
DB 54 MNSYISNVNIEGSGPOTTSNVIWPTGSHSPLVVRGQSHIYAPPOKSPWFAKSI 113
|||||

QY 61 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 120
|||||
DB 114 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 173
|||||

QY 114 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 173
|||||

QY 121 KESLGHNDYTCATNCTTTRKPKSQAVTRIPQVEVVMKQSSPFPQYPLVPP 180
|||||
DB 174 KESLGHNDYTCATNCTTTRKPKSQAVTRIPQVEVVMKQSSPFPQYPLVPP 233
|||||

QY 181 SADEPHGCAKAKRPSGSHAPVPELITLALSPPGVITLLEPFPVHLSPPSAFTFAS 240
|||||
DB 234 SADEPHGCAKAKRPSGSHAPVPELITLALSPPGVITLLEPFPVHLSPPSAFTFAS 292
|||||

QY 241 MMSLTKLADKEIVHMSAKKIPGVELSLFDQVRLLESQMEVLMGLMRSIDHCK 300
|||||
DB 294 MMSLTKLADKEIVHMSAKKIPGVELSLFDQVRLLESQMEVLMGLMRSIDHCK 353
|||||

QY 301 LIFAPNVLDPGKCVFGLLEFPMMLATTSRPFELIQRKEYLVKAMLLNSMYPL 360
|||||
DB 354 LIFAPNVLDPGKCVFGLLEFPMMLATTSRPFELIQRKEYLVKAMLLNSMYPL 413
|||||

QY 361 VIAQDASSPKIAHINAVTALVWVIAKSGSSQSSQSMPLANILMLSHVHANSCKM 420
|||||
DB 414 VIAQDASSPKIAHINAVTALVWVIAKSGSSQSSQSMPLANILMLSHVHANSCKM 473
|||||

QY 421 EHLL 424
|
DB 474 SHSL 477

RESULT 4
Q90ZEB
ID Q90ZEB PRELIMINARY: PRT: 486 AA.
AC Q90ZEB
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DB ESTROGEN RECEPTOR beta 2.
GN ERRETA2.
OS Gallinix jacchus (Common marmoset).
OC Eularchia: Metazoa: Chordata: Vertebrata: Euteleostomi.
OC Mammalia: Eularchia: Primates: Platyrrhini: Callitrichidae: Callithrix.
OX NCBI_TaxID=9483.
RN [1]
RP SEQUENCE FROM N.A.
KA Stobie G.A., Wilson J.A., Millar M.R., Macpherson S., Saunders P.L.:
RT "The estrogen receptor beta variant ERbeta.cx/ERbeta2 is expressed in
a wide range of tissues in both old and New World primates."
RL Submitted (JUN 2001) to the EMBL/GenBank/DBJ databases
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY
EMBL: AF93816; AAK71318.1;
DB InterPro: IPR000536; Hormone_rec_119.
DB InterPro: IPR001628; Zn1_C4steroid.
DB Pfam: PF00104; hormone_rec_1.
DB Pfam: PF00105; zf-C4_1.
DB ProDom: PD000035; Zn1_C4steroid_1.
DB PROSITE: PS00031; NUCLEAR RECEPTOR; UNKNOWN.1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SFGHVF 486 AA: 54344 MW: 48268573717777 CPE64:

Query Match 81.6%; Score 2058; DB 6; Length 486;
Best Local Similarity 90.8%; Pred. No. 9, 8e-182;
Matches 385: Conservative 17: Mismatches 22: Indels 0: Gaps 0;

QY 1 MNSYISNVNIEGSGPOTTSNVIWPTGSHSPLVVRGQSHIYAPPOKSPWFAKSI 60
|||||

DB 54 MNSYISNVNIEGSGPOTTSNVIWPTGSHSPLVVRGQSHIYAPPOKSPWFAKSI 113
|||||

QY 61 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 120
|||||

DB 114 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 173
|||||

QY 114 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 173
|||||

QY 121 KESLGHNDYTCATNCTTTRKPKSQAVTRIPQVEVVMKQSSPFPQYPLVPP 233
|||||

DB 174 KESLGHNDYTCATNCTTTRKPKSQAVTRIPQVEVVMKQSSPFPQYPLVPP 233
|||||

QY 181 SADEPHGCAKAKRPSGSHAPVPELITLALSPPGVITLLEPFPVHLSPPSAFTFAS 240
|||||

DB 234 SADEPHGCAKAKRPSGSHAPVPELITLALSPPGVITLLEPFPVHLSPPSAFTFAS 294
|||||

QY 241 MMSLTKLADKEIVHMSAKKIPGVELSLFDQVRLLESQMEVLMGLMRSIDHCK 300
|||||

DB 294 MMSLTKLADKEIVHMSAKKIPGVELSLFDQVRLLESQMEVLMGLMRSIDHCK 353
|||||

QY 301 LIFAPNVLDPGKCVFGLLEFPMMLATTSRPFELIQRKEYLVKAMLLNSMYPL 360
|||||

DB 354 LIFAPNVLDPGKCVFGLLEFPMMLATTSRPFELIQRKEYLVKAMLLNSMYPL 413
|||||

QY 361 VIAQDASSPKIAHINAVTALVWVIAKSGSSQSSQSMPLANILMLSHVHANSCKM 420
|||||

DB 414 VIAQDASSPKIAHINAVTALVWVIAKSGSSQSSQSMPLANILMLSHVHANSCKM 473
|||||

QY 421 EHLL 424
|
DB 474 SHSL 477

RESULT 4
Q90ZEB
ID Q90ZEB PRELIMINARY: PRT: 542 AA.
AC Q90ZEB
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DB ESTROGEN RECEPTOR beta.
GN Squalus acanthias (Spiny dogfish).
OS Eularchia: Metazoa: Chordata: Vertebrata: Chondrichthyes:
OC Elasmobranchii: Squala: Squalidae: Squalus.
OX NCBI_TaxID=7797.
RN [1]
RP SEQUENCE FROM N.A.
KA Sikora R., Bellard G.V.:
RT "Molecular cloning and stage-related expression of an Estrogen
Receptor Beta Subtype During Spermatogenesis in the Dogfish Shark
(Squalus Acanthias) Testis."
RL Submitted (MAY 1999) to the EMBL/GenBank/DBJ databases
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY
EMBL: AF147746; AAK57823.1;
DB InterPro: IPR000536; Hormone_rec_119.
DB InterPro: IPR001628; Zn1_C4steroid.
DB Pfam: PF00104; hormone_rec_1.
DB Pfam: PF00105; zf-C4_1.
DB ProDom: PD000035; Zn1_C4steroid_1.
DB PROSITE: PS00031; NUCLEAR RECEPTOR; UNKNOWN.1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 542 AA: 61285 MW: 5592929697180037 CPE64:

Query Match 61.8%; Score 1558.5; DB 13; Length 542;
Best Local Similarity 63.1%; Pred. No. 1, 7e-135;
Matches 294: Conservative 73: Mismatches 86: Indels 13: Gaps 5;

QY 1 MNSYISNVNIEGSGPOTTSNVIWPTGSHSPLVVRGQSHIYAPPOKSPWFAKSI 60
|||||

DB 77 LGSYMSDASGPGTIVROSLSPSYWSSGCHVSPITLHQQPIYAEAPKSPWIDLRSG 136
|||||

QY 61 FHILPVNPTTLPKVSQNPASVTPGSGPVAHFGVAVGASVYAGVSTGVYAF 120
|||||

```

DB 147 DQHL LNREMLKKKKIAPV--GSSVNVGVSRDAHFAVNDPAGVGVGWSVTEGKAF 194
CY 121 KPSGTHNVYGVATNAG--TQNKPKSVACPLKRYEYVWVCKSGRREPEVYKIVBGR 190
DB 194 KRSVGNNAVYGVATNAG--TQNKPKSVACPLKRYEYVWVCKSGRREPEVYKIVBGR 254
CY 181 SAQDPAIDVACAKRNSG GHAPVRRPDLALSPQGVTLTLEAPPPV--ISRPAPPE 238
DB 254 LKSSQVQJMSKRCRQISDTNLSIDCLMELAFERKLSSLEAFENPVNSLNHKKPTE 314
CY 239 ASNMSTETKADKEIVYMISMAKRTGVETLSLFGVRLLESTWREVMIMMIMWSIDHP 298
DB 414 VSMNSTETKADKEIVYMISMAKRTGVETLSLFGVRLLESTWREVMIMMIMWSIDHP 374
CY 299 GKLFAPDLVDRDECKVEGTLETFDMLATTSRPRELKQREYVAVKAMILLNSNMV 358
DB 474 GKLFAPDLVDRDECKVEGTLETFDMLATTSRPRELKQREYVAVKAMILLNSNMV 434
CY 459 PLVATGVANSSKRLAHNLNAVDAVWVIAKSGISSQGSRLANLMLSHVHNSNR 418
DB 434 FRSQVTEHENEKREKLRHLPDPLTWMSKSGTTPQQAIRKALMLSHVHNSNR 494
CY 419 GMRDYSMKRKNVYVYDLEMLNANHLKGRKSSITGSSCPAD 464
DB 494 GMRDYSMKRKNVYVYDLEMLNANHL 541

```

RESULT 5

```

DB 09RSM7 PRELIMINARY: PRT: 554 AA.
AC 09RSM7:
DB 01 JUN-2001 (TFEMBLrel_19, created)
DB 01 JUN-2001 (TFEMBLrel_19, last sequence update)
DB 01 JUN-2002 (TFEMBLrel_21, last annotation update)
DB ZEPER beta42 protein.
DB ZEPER BETA2.
DB Brachydanio rerio (zebrafish) (zebra danio).
DB Eukaryotic: Metazoa: Chordata: Vertebrata: Euteleostomi:
DB Actinopterygii: Neopterygii: Teleostei: Ostariophysi: Cypriniformes:
DB Cyprinidae: Danio.
DB NCBI Taxid 7955.
DB 111
DB SEQUENCE FROM N.A.
DB Mount A., Poliquini E., Andrade L., Blaise O., Landel V., Kahn G.
DB Jockel P.
DB "Binding characteristics, transactivation properties and central
DB expression of three estrogen receptor subtypes in zebrafish."
DB Submitted (SEP-2001) to the EMBL/Genbank/DDB databases.
DB 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DB 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DB EMBL: A1414567; CAC94849.1; ...
DB InterPro: IPR000546; hormone_rec_119.
DB InterPro: IPR001628; ZnF_C4steroi.
DB Pfam: PF00104; hormone_rec_1.
DB PRINTS: PR00105; Z1-C4_1.
DB PRODOM: PD000045; ZnF_C4steroi.1.
DB PROSITE: PS00041; NUCLEAR RECEPTOR; DNKNMNL.
DB DNA binding: Nuclear protein; Receptor; Transcription regulation;
DB Zinc finger.
DB SEQUENCE 554 AA; 64323 MW; 84C0DF667AC0F5ED CRR64;

```

```

Query Match 54.48; Score 1469; DB 13; Length 554;
Post local Similarity 58.28; Prod. No. 60-118;
Matches 270; Conservative 66; Mismatches 108; Indels 20; Gaps 7;
CY 18 KQTSNVNVPVTCNLSPLVNHQLSLVAELCKSNWFAVPSIETPLVNPPELKKVSG 77
DB 83 KQTSNVNVPVTCNLSPLVNHQLSLVAELCKSNWFAVPSIETPLVNPPELKKVSG 142
CY 78 NQASVGVGSRKDAHFAVNDPAGVGVGWSVTEGKAFKRSIGJNHYGVATNAG 147
DB 143 TEETS--VSLKRAKIMHFAVNDPAGVGVGWSVTEGKAFKRSIGJNHYGVATNAG 200

```

```

CY 148 CTIKNKRSVQVATLKKCYEYVWVCKSGRREPEVYKIVBGRKASDAV--LHACAKRS 195
DB 201 CTIKNKRSVQVATLKKCYEYVWVCKSGRREPEVYKIVBGRKASDAV--LHACAKRS 256
CY 196 GHAPVRRPDLALSPQGVTLTLEAPPPV--ISRPAPPE 245
DB 257 GPRSQELKSTPRPLSNVAVKISLSPEDLSIRMEAPPEVYKIMKRPPEAVVMSL 316
CY 246 LKALHPLVMSMAKRTGVETLSLFGVRLLESTWREVMIMMIMWSIDHPGKLFAP 305
DB 317 TNLADKRLVMSMAKRTGVETLSLFGVRLLESTWREVMIMMIMWSIDHPGKLFAP 376
CY 306 DVLVDRDECKVEGTLETFDMLATTSRPRELKQREYVAVKAMILLNSNMV YELVAT 364
DB 377 DLSLSPDSSCVQGVETLETFDMLATTSRPRELKQREYVAVKAMILLNSNMV TSSSG 416
CY 465 QDAISSKRLAHNLNAVDAVWVIAKSGISSQGSRLANLMLSHVHNSNRKREHL 424
DB 447 EDLQSRKSLCLDSVTDALVWVIAKSGISSQGSRLANLMLSHVHNSNRKREHL 496
CY 425 NMKKNVYVYDLEMLNANHLKGRKSSITGSSCPAD ERSK 467
DB 497 GMRDYSMKRKNVYVYDLEMLNANHLIMHNSRLSHVHNSNRKREHL 540

```

RESULT 6

```

DB 09RSM7 PRELIMINARY: PRT: 554 AA.
AC 09RSM7:
DB 01 JUN-2001 (TFEMBLrel_17, created)
DB 01 JUN-2001 (TFEMBLrel_17, last sequence update)
DB 01 JUN-2002 (TFEMBLrel_21, last annotation update)
DB Estrogen receptor beta4.
DB Brachydanio rerio (zebrafish) (zebra danio).
DB Eukaryotic: Metazoa: Chordata: Vertebrata: Euteleostomi:
DB Actinopterygii: Neopterygii: Teleostei: Ostariophysi: Cypriniformes:
DB Cyprinidae: Danio.
DB NCBI Taxid 7955.
DB 111
DB SEQUENCE FROM N.A.
DB Tissue-Liver.
DB Akten B., Kishida M., Gullard G.V.
DB "Estrogen receptor cDNAs in zebrafish."
DB Submitted (FEB-2001) to the EMBL/Genbank/DDB databases.
DB 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DB 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DB EMBL: AF49414; AKK5742.1; ...
DB HSPY: P03372; IERR.
DB InterPro: IPR000546; hormone_rec_119.
DB InterPro: IPR001724; Steroid_receptor.
DB InterPro: IPR001628; ZnF_C4steroi.
DB Pfam: PF00104; hormone_rec_1.
DB PRINTS: PR00047; STEROIDHORMONER.
DB PRODOM: PD000047; STEROIDHORMONER.
DB PROSITE: PS00045; ZnF_C4steroi.1.
DB SMART: SM00409; HELL_1.
DB SMART: SM00499; ZnF_C4_1.
DB PROSITE: PS00041; NUCLEAR RECEPTOR; 1.
DB DNA binding: Nuclear protein; Receptor; Transcription regulation;
DB Zinc finger.
DB SEQUENCE 554 AA; 62210 MW; A3BDC9B8E10HE1D CRR64;

```

```

Query Match 54.98; Score 1360; DB 13; Length 554;
Post local Similarity 57.88; Prod. No. 4-10-117;
Matches 268; Conservative 66; Mismatches 110; Indels 20; Gaps 7;
CY 18 KQTSNVNVPVTCNLSPLVNHQLSLVAELCKSNWFAVPSIETPLVNPPELKKVSG 77
DB 83 KQTSNVNVPVTCNLSPLVNHQLSLVAELCKSNWFAVPSIETPLVNPPELKKVSG 142
CY 78 NQASVGVGSRKDAHFAVNDPAGVGVGWSVTEGKAFKRSIGJNHYGVATNAG 147
DB 143 TEETS--VSLKRAKIMHFAVNDPAGVGVGWSVTEGKAFKRSIGJNHYGVATNAG 200

```

```

Db 143 TFEETS--VSLRGKAIHMYCAVCSFYASGYHCWSECECKAPFKKSTGCHNDYTCGPATNG 200
      148 CTIDKRRPSQQA*RIKCYEYGMKQSPREPOYRLVPRQSSADPD--LHCACAKARS 195
      201 CTIDKRRPSQQA*RIKCYEYGMKQSPREPOYRLVPRQSSADPD--LHCACAKARS 256
      196 GCHAPYVR-----ELLIDALSPQIVLITLAEPRHV-LISRPASPTFASMMNL 245
      257 GPRSEIKSLPRPLSNFVPRISLSPEELISPMIAEPEPEIYIMKMKPTEEAIVMNL 416
      246 TKLAKREIVHIMSWARKIPFVELSLFQVNPVLPESQWEMVLMKIMWSPTEPCLEIAP 405
      417 FNADKELVHIMSWARKIPFVELSLFQVNPVLPESQWEMVLMKIMWSPTEPCLEIAP 476
      406 DLVLDRECKVYGGILEIFDMLLATTSPREELKIQKEYLVYKAMITLNSM-YPLVYAT 464
      477 DLCTSPRESCVGVLFIFPMILATSPREELKIQKEYLVYKAMITLNSM-YPLVYAT 436
      365 GPDASSPKIAHLINAVTDALVWYIAKSGISSQGSRIANILMI SHVPHASNKGMHLL 424
      437 EDIASSRKILCVIINSTDAIWMATSTGTSQSPSTRIAHILMILSHVPHASNKGMHLL 494
      425 NMCKRVNVYVYDIIIFMNAHVLPGKSSITSSPSFA-EDSKS 467
      497 CMKMKKVVPIYDILLMLLAHIMSSPLSHSHAPAAHKKDKS 549

```

RESULT 7

```

Db 080HK9 PRELIMINARY: PRT: 565 AA.
AC 080HK9:
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Estrogen receptor beta.
GN ERB.
OS Paratichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygii; Acanthopterygii; Percomorphi; Pleurocentriformes;
OC Pleurocentriformes; Paratichthyidae; Paratichthys.
OX NCBI_TaxId=8255;
RN 11
RP SEQUENCE FROM N.A.
RA Kitano T., Sakimura N., Takamune K., Nagahama Y., Abe S.;
RT "Role of estrogen receptor in gonadal sex differentiation in Japanese
RT flounder (Paratichthys olivaceus).";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AH070630; BAB85623.1;
KW Receptor.
SQ SEQUENCE 565 AA; 64521 MW; 74CAB8F1426B087D CRC64.

```

Query Match

Best Local Similarity 53.7%; Score 1354; DB 13; Length 565;
Matches 276; Conservative 64; Mismatches 130; Indels 22; Gaps 10.

```

QY 2 NYSIPSNVTELEGNG-RQTTSPNVLPFPCHLS-PLVHR-QLSHVAFQSPQFAP 58
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 72 SYAPSS--IPGCFVHOSLSPLFWSHHNMIPMIIHSGVSGGQNPQPSQWCT 127
      59 SLHHLLPVNREIKRKYVSNRPAAPVTPDSSKRAHFAVSTAAVSTAAVSTAA 114
      128 PRGVLA-ANSKVPFR SQFSEGVVSSKRSGLHYCAVCHVAVCHVAVCHVAV 164
      119 FPRSTGGHNVGPATNCTIDKRRPSQQA*RIKCYEYGMKQSPREPOYRLVPRQ 176
      185 FFKRSILOGHNDYTCGPATNGTIDKRRPSQQA*RIKCYEYGMKQSPREPOYRLV 241
      177 PRGSALPFGHACAKAPSPQAPVPELITIAVSEPTVLTFAATGHW-LTCEVAT 200
      245 PRITLSSGSPASPKAITGPVVALMNPLOPAPITFQPIPMIAFPIVITAMTMSGP 404
      246 FTASMMNLTKLAKREIVHIMSWARKIPFVELSLFQVNPVLPESQWEMVLMKIM 296

```

```

Db 405 LLEAVNMSSLTHLAKREIVHIMSWARKIPFVELSLFQVNPVLPESQWEMVLMKIM 464
      296 DHEKILFAVGLVLDRECKVYGGILEIFDMLLATTSPREELKIQKEYLVYKAMIT 455
      465 DHQCKLFSPLSLSSSEGSQVGPSEIFDMLLATSKREELKIQKEYLVYKAMIT 424
      356 SM-YPLVYATQDASSRKIAHLINAVTDALVWYIAKSGISSQGSRIANILMI SH 414
      425 NMCLSSSEGSSELSHRSKILLIDVAVTDALVWYIAKSGISSQGSRIANILMI SH 484
      415 ASKGMHILNMCKRVNVYVYDIIIFMNAHVLPGKSSITSSPSFA-EDSKS 466
      485 VSNQMDHILCKMKKVVPIYDILLMLLAHIMSSPLSHSHAPAAHKKDKS 544
      467 SKRASON 473
      545 SONSSSN 551

```

RESULT 8

```

Db 090MS9 PRELIMINARY: PRT: 592 AA.
AC 090MS9:
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE ZHR-beta1 protein.
GN ZPER-BETA1.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN 11
RP SEQUENCE FROM N.A.
RA Menout A., Pellegrini E., Anglade I., Haise O., Landel V., Kah O.;
RA Pakdel F.;
RT "Blinding characteristics, transactivation properties and central
RT expression of three estrogen receptor subtypes in zebrafish.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AJ14566; GAC93848.1;
DR InterPro: IPR000345; Cyt_c_heme_bind.
DR InterPro: IPR000546; Hormone_rec_119.
DR InterPro: IPR001628; Zn_f_4steroid.
DR Pfam: Pf00104; hormone_rec_1.
DR Pfam: Pf00105; zf-C4_1.
DR ProDom: PD00035; Zn_f_4steroid_1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKN:UN.1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; UNKN:UN.1.
KW DNA binding, Nuclear protein, Receptor, Transcription regulation,
KW Zinc-finger.
SQ SEQUENCE 592 AA; 66046 MW; 7F05218E2C03F1B8 CRC64.

```

Query Match

Best Local Similarity 53.7%; Score 1354; DB 13; Length 592;
Matches 281; Conservative 63; Mismatches 98; Indels 54; Gaps 12.

```

QY 27 SAWEWLPDQ---SGLVVRQD SHVYATPSSWTAWSLPLGQVMSPTF-- 72
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 72 SPILLNWF-PSHMDQPTAVSLKSH-----SAAEAKI--HLSQSSVLSITKILG 141
      73 KVSNNCAAGVIGCKKRAHFAVSSVASYGVGWS/GDYVATFKRSIGCHNDYTC 142
      142 GCFHNMHNSA-SWVAGQHPGAVHCVANSYHGVWS-PGAKPRSTGCHNDYTC 201
      103 FALNGLIIGNREFS*ACVLEFEVSWVPTZEEELGCVLVVPPG---SAPFOL 186
      202 FALNGLIIGNREFS*ACVLEFEVSWVPTZEEELGCVLVVPPG---SAPFOL 261
      187 HCAGKAKR-----SGHAPVRELLIDALSPQIVLITLAEPRHV-LISRP 242

```


284 EVIMMGLAMRSIDHCKLIFADVLVDHCKCVETLEIFIMLATSSEPELKIARE 343
 470 ELLIVH IWRSTIPPGQHEFAPNI ILCPDANVPMLEPEMLLVSPPELIPER 429
 444 YLCVAKMLLNSSM-YPIVATATODSSKLAHLNAVTDALVWIAKSGISQOOSRI 402
 430 YVIAKMLLNSGVEFCISNSAGROTNVQLIQDILEKMDAGSTTHIFASPIQHSRI 489
 403 ANILMLSHVRIANSKMEHLNKKCNVIVYDLEMLNHLVLRGKSSITGSESP 461
 490 SOLLLLSOIRHLSNKGIELHNSMKRKVPIEYDLEMLDASHION-----TGLRSP 543

RESULT 19

Q990J3 PRELIMINARY: PRI: 458 AA.

AC Q990J3
 DT 01-MAR-2001 (Tremblrel, 16, Created)
 DT 01-MAR-2001 (Tremblrel, 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
 DE Estrogen receptor (Fragment).
 GN ER.
 OS Eukaryotes: Chordata: Craniata: Vertebrata: Euteleostomi:
 OC Actinopterygii: Neopterygii: Teleostei: Psectrostei, Neoteleostei,
 AC Acanthomorpha: Acanthopterygii: Percormorpha: Perciformes: Labroidae:
 OC Labridae: Halichoeres.
 OX NCBI_TaxID:147232;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kim S.J., Koi G., Takemura A., Nakamura M.
 RT Partial sequence and expression of androgen and estrogen receptor
 RI genes in the proteoglycan wrasse, halichoeres trimaculatus.*
 RL Submitted (08-2000) to the FMB/GenBank/EMBL databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC EMBL: AF349413; AAK16741.1;
 DR HSSP: P03372; IERR.
 DR InterPro: IPR000536; Hormone_rec_119.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; ZnF_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; zf-C4_1.
 DR PRINTS: PR00398; STEROHORMNER.
 DR PRINTS: PR00047; STROLDPNSGR.
 DR ProDom: PD000035; ZnF_C4steroid_1.
 DR SMART: SM00430; HOL1_1.
 DR SMART: SM00399; ZnF_C4_1.
 DR PROSITE: PS00041; NUCLEAR_RECEPTOR_1.
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 FT NON_TER
 ST SEQUENCE 458 AA: 51319 MM: 0194823321119777 C9964;

Query Match

Best local similarity: 43.6%; Score 1099; DB 13; Length 458;
 Matches 210; Conservative 64; Mismatches 85; Indels 40; Gaps 10;

QY 91 RIAHFAVNSVASTYHYVWSYPTKAF+PSLGSHNYTPDANQOTLKKPKSVA 150
 23 KEMRYAVNSVASTYHYVWSYPTKAF+PSLGSHNYTPDANQOTLKKPKSVA 82
 QY 151 GPRKRYVNSVASTYHYVWSYPTKAF+PSLGSHNYTPDANQOTLKKPKSVA 197
 DB 83 CHIRKRYVNSVASTYHYVWSYPTKAF+PSLGSHNYTPDANQOTLKKPKSVA 146
 QY 198 HARVPRFIIFAIS--PQVVIITIFAPPHVLSPP--SAFTFASMMMLFRLAREL 254
 DB 137 NATIVEENQWIMSPDQVLLDQAFITP--ILYSROKLSRPTVIMMLITLSMADREL 195
 QY 254 VHIISAKKIPQVEISTFQVPILESQMEVLMGMIMPSIDHCKLIFADVLVDHCK 313

DB 196 VHIISAKKIPQVEISTFQVPILESQMEVLMGMIMPSIDHCKLIFADVLVDHCK 255
 QY 414 GRCVETLEIFIMLATSSEPELKIAREYIVAKMLLNSMYPVATATOD SSPEK 472
 DB 256 GRCVETLEIFIMLATSSEPELKIAREYIVAKMLLNSMYPVATATOD SSPEK 416
 QY 473 LAHLNAVTDALVWIAKSGISQOOSMIALMLSHVRIANSKMEHLNKKCNVIV 432
 DB 416 VQMDITITDALIHHSQSAHQSPQACITITISHIPMSKMEPHVSKMCPKVK 476
 QY 433 PYVDLEMLNHLVLRGKSSITGSESPADSKRSGSONQS 476
 DB 476 PYVDLEMLNHLVLRGKSSITGSESPADSKRSGSONQS 424

RESULT 20

Q98SM8 PRELIMINARY: PRI: 391 AA.

AC Q98SM8
 DT 01-JUN-2001 (Tremblrel, 17, Created)
 DT 01-JUN-2001 (Tremblrel, 17, Last sequence update)
 DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
 DE Estrogen receptor gamma (Fragment).
 GN Brachydanio rerio (zebrafish) (Zebra danio).
 OS Eukaryotes: Chordata: Craniata: Vertebrata: Euteleostomi:
 OC Actinopterygii: Neopterygii: Teleostei: Ostariophysi: Cypriniformes:
 OC Cyprinidae: Danio.
 OX NCBI_TaxID:7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Arden B., Kishida M., Gallard G.V.
 RT "Estrogen receptor cDNAs in zebrafish.*"
 RL Submitted (Feb-2001) to the EMBL/GenBank/EMBL databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC EMBL: AF349413; AAK16741.1;
 DR HSSP: P03372; IHCO.
 DR InterPro: IPR000536; Hormone_rec_119.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; ZnF_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; zf-C4_1.
 DR PRINTS: PR00398; STEROHORMNER.
 DR ProDom: PD000035; ZnF_C4steroid_1.
 DR SMART: SM00430; HOL1_1.
 DR SMART: SM00399; ZnF_C4_1.
 DR PROSITE: PS00041; NUCLEAR_RECEPTOR_1.
 KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
 KW Zinc-finger.
 FT NON_TER
 ST SEQUENCE 391 AA: 45E9612666A816E91 C9964;

Query Match

Best local similarity: 42.1%; Score 1061.5; DB 13; Length 391;
 Matches 219; Conservative 55; Mismatches 67; Indels 45; Gaps 7;

QY 133 PATNOSTDKSPKSGVATPKCYVNVWVCEPPPVVYPRKOR-----SAFQEL 186
 DB 1 PATNOSTDKSPKSGVATPKCYVNVWVCEPPPVVYPRKOR-----SAFQEL 60
 QY 187 HCAKAKR-----SGHAPVRELLDASPEQVLTLEAPPHVILSKP 242
 DB 61 GLEPGSQOHFPPISQSLHPLPSGPA--EGPGL--NVSPELVLTLEAPPHVILSKP 116
 QY 233 -SAFTFASMMMLFRLARELIVAKMLLNSMYPVATATOD SSPELKIAREYIVAK 291
 DB 117 VKKPYTFASSMMMLFRLARELIVAKMLLNSMYPVATATOD SSPELKIAREYIVAK 176
 QY 292 WPSIDPDKITFADVLVDHCKCVETLEIFIMLATSSEPELKIAREYIVAKML 351
 DB 177 WPSIDPDKITFADVLVDHCKCVETLEIFIMLATSSEPELKIAREYIVAKML 246
 QY 352 LLNSMYPVATATOD SSPELKIAREYIVAKMLLNSMYPVATATOD SSPELKIAREYIVAK 410

RA Ye J., Yeh K., Fu Z., Gaveli, Zhan M., Chang C., Chen G., Chen G.,
 Zhang X.H., Zhang F.N., Zhou W., Zhou X., Zhu G., Zhu X., Smith B.,
 RA Chao B.A., Myers E.W., Rubin G.M., Wooten J., et al. Zhu X., Smith B.,
 RA "The genome sequence of *Drosophila melanogaster*."
 RA Science 287:2185-2194 (2000).
 RA [2]
 RA SEQUENCE FROM N.A.
 RA STRAIN BEKLEY.
 RA Stojanovic M., Hrochova P., Hong L., Ardayan A., Carlson J.,
 RA Clump M., Chavez C., Dorset V., Fattori D., Friso E., George K.,
 RA Gonzalez M., Gauthier J., Li P., Liao G., Miranda A., Mungall G.J.,
 RA Munoz J., Pauley J., Paragas V., Park S., Phoonuwong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Colniker S.,
 RA Submitted (Aug. 2001) to the EMBL/Genbank/DBM databases.
 RA [3]
 RA SEQUENCE FROM N.A.
 RA Lehoucq J., Wherry S., Talmadge C., Davis J.R., Kiss L., Fudy J.D.,
 RA Chao B.A., Smoqi J.,
 RA "Identification and characterization of the *Drosophila* ortholog of
 RA vertebrate estrogen-related receptor gene and its product."
 RA Submitted (MAR 2001) to the EMBL/Genbank/DBM databases.
 RA [4]
 RA SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 RA EMBL: AF504556; AAF50474.1;
 RA EMBL: AF505162; AAK94056.1;
 RA EMBL: AF594201; AAL37554.1;
 RA BSSB: J04372; HCG.
 RA EMBL: F600035449; ERR.
 RA InterPro: IPR000546; Hormone_rec_119.
 RA InterPro: IPR001724; Steroid_receptor.
 RA InterPro: IPR001628; Zn1_GAsteroid.
 RA Pfam: PF00104; hormone_rec_1.
 RA Pfam: PF00105; z1_c4_1.
 RA PRINTS: PR00398; STRODERMONER.
 RA PRINTS: PR00047; STRODERMONER.
 RA Prodom: PR000045; Zn1_GAsteroid.1.
 RA SMART: SM00430; H01.1.
 RA SMART: SM00399; Z06_C4_1.
 RA PROSITE: PS00041; NUCLEAR_RECEPTOR.1.
 RA DNA binding: Nuclear protein receptor; Transcription regulation;
 RA Zinc finger.
 RA Sequence: 494 AA; 5429 MW; 2548AA078A411 CR664;
 Query Match: 20.7%; Score 522; DB %; Length 484;
 Host Local Similarity: 41.5%; Prod. No. 1,1e+49;
 Matches: 155; Conservative: 77; Mismatches: 174; Indels: 86; Gaps: 18;
 DB 4 YSIPSVNTLEGGPQDTSPVNIWPTGHSPLVNHQSLHVAEPKSPWPARSLER 62
 DB 22 FSPSSKSTATGSGTNGIKSSPSV-----SP-----EKL-----CSSTTSLSCHIDNV 64
 DB 63 TLVNRRETLKRK-----VSQNRG--ASPVTPGSKRD--AHFVAVSDYASG 105
 DB 65 SLNSKIDSLKSGTSGNGGSGGSGTSGNATNASAGCSGVDRDELRLGLVGVYASG 124
 DB 106 YHGVMSGECKAKFEKRSLOGINDYTPATNOCTIDKMKRSQACRLKRYEYGMVKG 165
 DB 125 FHGVASCEACAKFEKFTLQGNIEYTPANNPEELNKRKKRAGQAKRQKTLMDKKEG 184
 DB 166 SR--KRGYRLVRKQ--KSAIDRLRACAKKRSQADAKRRELLALDAISPEQVLTILE 221
 DB 185 VRLDKVRGQKQKRRKRVNSVYQMDLYQSNITSLIVKILE-VLNSVEYDAL--SVQ 240
 DB 222 AEPPV--LISPSAPTEFA--MMMS-----LKLAKRE 252
 DB 241 TTPPVHTTSTINDVASSSSGSIKLESSVVTNCTGTPONNNNNINPEILSVLSIDYKE 300
 DB 253 LVHMSWARKLPCEVLSLFDVRLKESQWMEVLMKIMKWSIDPKELIFAPRLVLRD 312
 DB 401 LVSVIGWAKQIPCEPILPLINDQMKLLQVSWAEITLQDTPRSIDPKNKICFATVIMWDEH 360
 DB 413 CKCCEGIIEDMLATTSPPRELKQKEEYLVKAMIDLNSMSPVLTATQIADSSR 371

DB 414 LAKT--SYLEFYH-VGAGMDLSEKELLYLALALANLQI 414
 DB 422 TLALHLLNAVTRLVWVIAFSSISGSGSMRANLMI-SVYPPASNV-AMHLLMMKVN 430
 DB 414 AFRGTTNSINNVYLLNHSVASHQQ 467
 DB 431 VWPVDELEML 442
 DB 468 VTKMKELFEML 479
 BEST1 26
 DB 10 ORMS79 PRELIMINARY; PRT: 496 AA.
 DB 01 MAR-2002 (TEMBLrel, 20, created)
 DB 01 MAR-2002 (TEMBLrel, 20, last sequence update)
 DB 01 JUN-2002 (TEMBLrel, 21, last annotation update)
 DB Drosophila melanogaster receptor splice variant.
 DB Drosophila melanogaster (Fruit fly).
 DB Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 DB Pterygota; Neoptera; endopterygota; Diptera; Hymenoptera; Muscomorpha;
 DB Ephydroidea; Drosophilidae; Drosophila.
 DB NCBI_TaxID=7227;
 DB [1]
 DB SEQUENCE FROM N.A.
 DB Lehoucq J., Wherry S., Talmadge C., Davis J.R., Fudy J.D., Chao B.A.,
 RA Kiss L., Smoqi J.,
 RA "Identification and characterization of the *Drosophila* ortholog of
 RA vertebrate estrogen-related receptor gene and its product."
 RA Submitted (MAR 2001) to the EMBL/Genbank/DBM databases.
 RA EMBL: AF594211; AAL37554.1;
 RA InterPro: IPR000546; Hormone_rec_119.
 RA InterPro: IPR001724; Steroid_receptor.
 RA InterPro: IPR001628; Zn1_GAsteroid.
 RA Pfam: PF00104; hormone_rec_1.
 RA Pfam: PF00105; z1_c4_1.
 RA PRINTS: PR00398; STRODERMONER.
 RA PRINTS: PR00047; STRODERMONER.
 RA Prodom: PR000045; Zn1_GAsteroid.1.
 RA SMART: SM00430; H01.1.
 RA SMART: SM00399; Z06_C4_1.
 RA PROSITE: PS00041; NUCLEAR_RECEPTOR; UNKNOWN.1.
 RA Receptor.
 RA Sequence: 496 AA; 54529 MW; 5067FSA8E9D3842 CR664;
 Query Match: 20.7%; Score 522; DB %; Length 486;
 Host Local Similarity: 41.5%; Prod. No. 1,1e+49;
 Matches: 155; Conservative: 77; Mismatches: 174; Indels: 86; Gaps: 18;
 DB 4 YSIPSVNTLEGGPQDTSPVNIWPTGHSPLVNHQSLHVAEPKSPWPARSLER 62
 DB 34 FSPSSKSTATGSGTNGIKSSPSV-----SP-----EKL-----CSSTTSLSCHIDNV 76
 DB 63 TLVNRRETLKRK-----VSQNRG--ASPVTPGSKRD--AHFVAVSDYASG 105
 DB 77 SLNSKIDSLKSGTSGNGGSGGSGTSGNATNASAGCSGVDRDELRLGLVGVYASG 136
 DB 106 YHGVMSGECKAKFEKRSLOGINDYTPATNOCTIDKMKRSQACRLKRYEYGMVKG 165
 DB 137 FHGVASCEACAKFEKFTLQGNIEYTPANNPEELNKRKKRAGQAKRQKTLMDKKEG 196
 DB 166 SR--KRGYRLVRKQ--KSAIDRLRACAKKRSQADAKRRELLALDAISPEQVLTILE 221
 DB 197 VRLDKVRGQKQKRRKRVNSVYQMDLYQSNITSLIVKILE-VLNSVEYDAL--SVQ 252
 DB 222 AEPPV--LISPSAPTEFA--MMMS-----LKLAKRE 252
 DB 253 LVHMSWARKLPCEVLSLFDVRLKESQWMEVLMKIMKWSIDPKELIFAPRLVLRD 312
 DB 401 LVSVIGWAKQIPCEPILPLINDQMKLLQVSWAEITLQDTPRSIDPKNKICFATVIMWDEH 362


```

QY 413 PCK-CVPGHLETFMMLATTSPPPEKIQHEEYLVKAMILLNSSMPLVATQVANSR 371
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 373 LAEC-CVTFEYFVWQAPMEPLSPPEEYLLKALLANCI-----LLDPSSELR 425
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 422 FVNH-LENAVDAIWWAKSQTSSQCSMSANIMLSHVEASNNMFLIMPECFN 430
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 426 AFRDTLNSINDVYVLLKRSSAVSHOO-----LILLPSLRQADILIRRWKTADE 479
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 431 VVIVYDLLEML 442
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 480 VITMKKLEML 491
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 27
046402 PRELIMINARY: PRT: 159 AA.
ID 046402
AC 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DI 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
DE Estrogen receptor alpha (Fragment).
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus;
   NCBI_TaxID=9824;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsu S., Winters T.A.;
RT "Swine estrogen receptor alpha hormone-binding domain, exon V to
   VIL1."
RL Submitted (NCV-1967) to the EMBL/Genbank/DDBJ databases.
DR EMBL: AF035775; AAB8509.1; -
DR HSSP: P03372; 3ERD.
DR InterPro: IPR030546; Hormone-rec_114
DR InterPro: IPR001723; Steroidm_receptor.
DR Pfam: PF00104; hormone_rec_1
DR PRINTS: PR00398; STRDH-RMNER.
DR SMART: SM00430; HOL1.
KW Receptor.
FT NON_TER 1 159
ET NON_TER 159
SQ SQUONPQ 159 AA: 14464 MW: 8344053E8E2A98 Dk554.

Query Match 20.5%; Score 517.5; DB 6; Length 159;
Post Local Similarity 59.5%; Pred. No. 6,7e-40;
Matches 94; Conservative 39; Mismatches 24; Indels 1; Gaps 1;

QY 285 VIKMMLWPSIDPCKEIPAPPIVTPDPEKCVPGHLETFMMLATTSPPPEKIQHEEY 344
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 2 1IMIGTVPSEHNGKLEFAPNLLDPNCKCVESWVEFDMILATSSPERMNIQCEHF 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 345 LCVKAMILLNSSMPLVATQVANSR-GLQSSPKLAHLNNAVITALVWVIAKSSISQCSMELA 403
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 62 VCKSTLTLSNVVTFPSSTKSEFPRKHHPVIRKTPVYIIMAKAVQITQCVGHPVA 121
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 404 NLMLSHVHASNKFMFHLINMKKVVNVYVLLLEML 441
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 11: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 122 QILLLSHFHMSKNKIMENLYNKKCNVYVYVLLLEML 159
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 28
045J01 PRELIMINARY: PRT: 249 AA.
ID 045J01
AC 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DI 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Estrogen protein (Fragment).
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis;
   NCBI_TaxID=9824;

```

```

   NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-EPIDIDYMIS;
RA Kaeppler-Hanno K.G.;
RT Submitted (JUN 2003) to the EMBL/Genbank/DDBJ databases.
QY 121
1b 121
QY 89 SKRDNHCAVSDYASGHYGVSCGCKAFKRSIQGHNDYICATINQTTIDKRRKSC 148
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 79 SAKTIPYCAVGNVASYHYGVSCSGKAFKRSIQGHNDYICATINQTTIDKRRKSC 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 149 GACVPCVGVVWVKSPPGPGVPLVPPQSAVPOHICACAPKPSGMAVYV----- 203
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 139 QACNLEKCYFVNMKATTFKDPKDPKMKRHKQLEFQ GPQNVVSSVSTSSLM 194
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 204 --RELLD-----ALSPDLVLLLEAPPHVLLS-RPSAPTEASMMKSL 246
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1b 194 FSHLEKHKRNHAGLIALQWVSAIILALHLYSDYISKEFSASMMKSL 248
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 29
041B05 PRELIMINARY: PRT: 710 AA.
ID 041B05
AC 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DI 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Progesterone receptor.
OS pr.
OC Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
   Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
   Anguilla
   NCBI_TaxID=7937;
QY 111
1b 111
QY 1150E-TESTIS;
RA Tissue-Testis;
RA Yoshikuni M., Yamauchi K., Nagahama Y.;
RT "Characterization of a testicular 17,20-dihydroxy-4-pregnen-3-one (a
   spermatogen inducing steroid in fish) receptor from a teleost,
   Japanese eel (Anguilla japonica)."
DI FEBS Lett. 465:12-17(2000).
OC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AB022075; BA089539.1; -
DR HSSP: P06536; JKGD.

```

```

108 InterPro: IPR000546; Hormone_rec_119;
109 InterPro: IPR001723; Steroid_receptor;
109 InterPro: IPR001628; Znf_C4stroid;
109 Pfam: PF00104; hormone_rec_1;
109 PRINTS: PR00498; STROHOMNER;
109 PRINTS: PR00047; STRODIFINER;
109 Prodom: PD000045; Znf_C4stroid; 1;
109 SMART: SM00430; HELL; 1;
109 SMART: SM00499; ZNF_C4; 1;
109 PROSITE: PS00041; NUCLEAR_RECEPTOR; 1;
109 DNA-binding: Nuclear protein; Receptor; Transcription regulation;
109 ZINC-finger;
109 SEQUENCE: 710 AA; 78819 MW; 0C8149H057A055AA CRR64;

Query Match: 20.0%; Score: 504; DB: 13; Length: 710;
Best Local Similarity: 41.4%; Prod. No. 6048;
Matches: 149; Conservative: 84; Mismatches: 145; Indels: 66; Gaps: 19;

108 43 LSPLEVVR - - - QLSHVAEKKSP - WCEAP - SEPTFPVN - - - - - PETLKPSVSN 78
109 271 LGLPVKSPSPKSAASA - IALSPKSPQSTSPSTPHLGLIYLSFALHSK - - - - - KYST 427
109 79 RCASPVTVGSKKRAHFAV - SDYASGVYGVWSTFCAKPKRSYGGHNYTPATNG 148
109 428 NAYSSYLAVLPGR - - - WVLGGPQASVCHYVLTGSGKVFPRKAVGHNNYLAQGRNR 384
109 139 TTRNPKSPQAPCPPEVYVAVKVGSGSPRPG - - - - - YVLVPGSSADPG 185
109 485 LVKTRKNTVA - RLKQYVAGMILGTRKIKKIALKAAGLQVALVANSITTRKSSGSSQ 444
109 186 LIRAGKAKSGGHAAPVRELLDALNSPOLVLTLEAPPHVLS - - - - - RESAPTEASMM 243
109 445 - - - - - ALMTATLPVREL - - - - - HLSYQ - - - - - LSVLESTPEVYVSGVNSQPMNMLN 493
109 244 STCKALKELVHMSAKF - - - - - EVLSLEFGVLES - - - - - KNEVIMMLMPSLPH 401
109 494 STNDA - - - - - RGLKRVNMSKSLGGRSLTINOMAL - - - - - IYVMSKSLMVSLSKMSQVNTSYL 553
109 402 LFAVDVLDDEK - - - - - VEGTLEEDMLA - - - - - TTSRRELKQHKREYLA - - - - - VKAMILLNSMY 458
109 554 YFALDGLINEYMK - - - - - RSPLEIDAMQVTFHGRANLQVTRKREFTMKVILLNTYVL 609
109 659 PLVATVDALSSKRLAHLLNAVTDALVWIAKSGISSVOOSMGLANLMLLSHVDAASNK 418
109 610 FGRKSQVDPKMG - - - - - NYHEIKAI - - - - - HLKRGVVA - - - - - RYVHLIKLMHMDLVAK 663
109 419 GMEHLMKK 428
109 664 - - - - - LHYG 668

RESULT 10
Q96EY4
AC Q96EY4 PRELIMINARY; ORF: 583 AA.
108 01-MAR-2001 (TEMBLrel: 16; Created)
109 01-MAR-2001 (TEMBLrel: 16; Last sequence update)
109 01-JUN-2002 (TEMBLrel: 21; Last annotation update)
109 Prosite: Prosite PR.
109 Xoonpus laevis (African clawed frog);
109 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
109 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
109 Xoonpus;
109 NCBI_TaxID: 8455;
109 8N
109 8P SEQUENCE FROM N.A.
109 8X MEDLINE: 20524046; PubMed: 11950156;
109 8A Bayat M., Booth R.A., Sheng Y., Lin X.;
109 8T "The classical progesterone receptor modulates Xoonpus oocyte
109 8U maturation through a non-genomic mechanism";
109 8L Proc. Natl. Acad. Sci. U.S.A. 97:12607-12612(2000);
109 8C 1. SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY);

```

```

109 1. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
109 EMBL: AY007198; AAC01366.1; -
109 BSSP: P06536; 1KCD;
109 InterPro: IPR000546; Hormone_rec_119;
109 InterPro: IPR000128; Progest_receptor;
109 InterPro: IPR001723; Steroid_receptor;
109 InterPro: IPR001628; Znf_C4stroid;
109 Pfam: PF00104; hormone_rec_1;
109 Pfam: PF02161; Prog_receptor_2;
109 PRINTS: PR00498; STROHOMNER;
109 PRINTS: PR00047; STRODIFINER;
109 Prodom: PD000045; Znf_C4stroid; 1;
109 SMART: SM00430; HELL; 1;
109 SMART: SM00499; ZNF_C4; 1;
109 PROSITE: PS00041; NUCLEAR_RECEPTOR; 1;
109 DNA-binding: Nuclear protein; Receptor; Transcription regulation;
109 ZINC-finger;
109 SEQUENCE: 583 AA; 66156 MW; 63BD4572C42B066 CRR64;

Query Match: 19.0%; Score: 478; DB: 13; Length: 583;
Best Local Similarity: 27.5%; Prod. No. 16045;
Matches: 146; Conservative: 88; Mismatches: 157; Indels: 114; Gaps: 19;

108 3 YSTPSVNTNLRSPRQSTSPNVLTPTGHLSPVYHROSLHVAEKKSPWFA - - - - - 57
109 136 FKLISHNTNDSGLTSTSAQTL - - - - - QHLSLN - - - - - GHVYTFQPTTRKREYTLQI 186
109 58 - - - - - RS - - - - - LEHTLVNRETLRKKSNNR - - - - - CASPVTVGSKKRAHFAV - - - - - SYVAGY 106
109 187 QLVVYVYTRSDRPERGQTPSPKMLDQK - - - - - - - - - - - CLGGPQASG 226
109 107 HGVWYVTVGKAFKPKSLGGHNYTPATNGTLDNKRNS - - - - - VAGRLKCFVAVKVGTS 166
109 227 HGVLTGSGKVFPRKALFGHNYLCAGRNDTVKTRKN - - - - - PSRLKQVAGAVLGR 286
109 167 KREKQYRLVGRKRSAD - - - - - - - - - - - QIRAKAKK - - - - - SGHAAPVRELLDALNSPOL 215
109 287 KPRKPG - - - - - RIKGRLEIDVVLSPPTSLPQVLLIKRISNSAQEIQ - - - - - FPGP - 346
109 216 VILLTFAEPHVLS - - - - - RPSAPTEASMMSTLKLQKELVHMSAKLPGVFLSPD 273
109 437 LLOTLDSPEVYVAGVYDTPPTPSALSSLNQLEKQV - - - - - VVWKSLSLQPRKRLDD 396
109 274 GVEILLSCMEVIMMLMPSLIDPG - - - - - KLFAPIVLDKPK - - - - - VQVLLPFL 424
109 497 QTLTQVMSLWVFAVMSVYGVHSGVYAPDILLNFGMKDSSFTYTLIS - - - - - 450
109 425 DMLATTSRRELKQHKREYLA - - - - - VKAMILLNSMYPLVATVDALSSK - - - - - LAHLLNAVTD 482
109 451 - - - - - RMQVLPQFKQVTFEELQMKALLNLI - - - - - LPRDLKSGVNFDMKSNVTFELAKVLS 507
109 483 ALWVIAKSGISSVOOSMGLANLMLLSHVDAASNKMDH - - - - - LNMKSKV - - - - - 431
109 508 - - - - - LHKGVYASSQVYQITKLM - - - - - - - - - - - ISMHELVKQHLVGLNTPQSRKLS 552
109 432 VVYVLDLLEMLNAVH 446
109 553 VEPPEMSEVLSMDL 567

RESULT 11
Q96I02
AC Q96I02 PRELIMINARY; ORF: 423 AA.
108 01-DEC-2001 (TEMBLrel: 19; Created)
109 01-DEC-2001 (TEMBLrel: 19; Last sequence update)
109 01-JUN-2002 (TEMBLrel: 21; Last annotation update)
109 Unknown (protein for MACH3401880) (Fragment);
109 Homo sapiens (Human);
109 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
109 Mammalia; Primates; Catarrhini; Hominoidea; Homo;
109 NCBI_TaxID: 9606;

```


[illegible]

QY	9	ABHCAVMSYASVAYHGWSTHKKKAFKRSLOCHINYOIPALINOUTHOKNRKSSQAR	15
DB	422	--PELTJOGHBASCHYGVLTGOSKVFYKFAVECHNOYIPALINODIVARKNPAVR	379
QY	153	LRKCYVWAKNSRREJG-----YKLVGHSALGOLHAKARKSSQADPAVR	203
DB	480	LRKCYVWAKNSRREJG-----YKLVGHSALGOLHAKARKSSQADPAVR	445
QY	204	ELIIDLSPGCVITLLEADPPHVLISPSA--PEFASMMMSKTRKAPFVHIMISAK	261
DB	436	EL-----GFTPO--ELSLTLEIEFEYVSQYADOTPEFPHLELNSINLECEKOLIMIVRMSK	490
QY	262	KTHPEVELSLPVVHLLSCGMEVIMMELMMSLHJBBCK--LIPADIVIDPEPCKVHG	319
DB	491	SLDGEFSLIMNOOHTLOYSKMSLWVSLDHRSSQVNTKEFLYFADHLLIDERRK----	546
QY	420	LELIFMLLA-----TTSKREELKLOHREYLVKAMLLNNSMPLVATQVADSSKRLAH	376
DB	547	NSPISLIDCAMMOIIPVAFINLOVTKKEFLCEKRVLLINT--VPLEGHSQAFDEMRBGY	604
QY	477	LNQVTLVAVWYLAKSSLSQSSQSMPLAMLLSHVHVASNKKMREMLINMK	428
DB	605	LRRLTKATQUTERKQVMASSQ--KEFLYLKLMIAHMLIVRK--VNLVY	647

RESULT	55		
CHROMID			
1D	COORDINATE	PRELIMINARY:	PRJ: 732 AA.
AC	COORDINATE		
DT	01-MAR-2001 (TEMPORAL_16, Created)		
DT	01-MAR-2001 (TEMPORAL_16, Last sequence update)		
DT	01-MAR-2002 (TEMPORAL_21, Last annotation update)		
1E	Progestrone receptor,		
EN	XPR_1.		
OS	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;		
OC	Xenopodidae; Xenopus.		
OX	NCBI_TaxID:8455;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEELINE 20570484; PubMed 1114187;		
RA	Tsian J., Kim S., Heilig F., Kuderman J.V.;		
R1	"Identification of XPR-1, a progesterone receptor regulated for Xenopus		
	oocyte activation."		
RL	Procyt. Natl. Acad. Sci. U.S.A. 97:14358-14363(2000).		
CC	1. SIMILARITY LOCATION: NUCLEAR (BY SIMILARITY).		
CC	1. SIMILARITY BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.		
EMBL	AF273435; AA042462.1; -		
1R	HSPF: P05345; 1r3D.		
1R	InterPro: IPR000546; Hormone_rec_114.		
1R	InterPro: IPR00128; Progest_receptor.		
1R	InterPro: IPR001723; Steroid_receptor.		
1R	InterPro: IPR01628; Znf_Cstteroid.		
1R	pfam: pf00104; hormone_rec_1.		
1R	pfam: pf02161; prog_receptor_1.		
1R	pfam: pf00105; z1-c4_1.		
1R	PRINTS: PR00498; STRDOMONR.		
1R	PRINTS: PR00647; STRIDFINDER.		
1R	ProDom: PD000045; Znf_Cstteroid_1.		
1R	SMART: SM00440; HELL_1.		
1R	SMART: SM00499; ZNF_C4_1.		
1R	PROSITE: PS00041; NUCLEAR_RECEPTOR_1.		
1R	DNB-binding: Nuclear protein receptor; Transcription regulator;		
1R	DNB-binding: Nuclear protein receptor; Transcription regulator;		


```

Db      84 GIKASGKHYGVSYSTKGGFEKRYVKKLSYVPPPTITIKRQBNQYGVYQK---- 140
QY      160 GNVKGGSPRRPGVAVRQPSADQLHQVQKAKSSGAPVPELLDALSPYAVETL 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      141 --LAKMKKFAVQEEQRTKDPADSEV-----ESTSGAP-----PEMLERI 181
QY      220 LEAPPPHVIISPPSAPTFASM----MMSITKIAKELVHIMSKAKKIPGVHLSIEVY 275
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      182 LEAF---LRVESQTIQLSESAGQDPVSSIQVAALRQLQVMAKHIPHEELPLEDKM 238
QY      276 PLESTVMEVIMMIMMESTIPPKETIAPPVVITPPQKVEFELIEPMLA-TSPF 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      239 VELKAMNFIILASHSRVSVPQVATPIVVPVPHAGCA-GVQALFVPELVIAKM 297
QY      335 PEIKIQHKRYGVKAMILLNSSMPVATVPANSSPKLA----HLLNAVTVAVWVIK 390
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      298 REMKMIRFELGLAVVLEPNFAKGLRTPSGRSGESVSALFEHQQYDP----- 349
QY      391 SGLSSQGSMLANIMLSHYHASNKMEHLLMKCKNVVPYDILLEMLNA 444
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      350 -----QPGKFAKILLRLPALRSGLKTEHLPFKLIDTPIIDFLSMLEA 396

```

RESULT 40

QRTSG6

ID QRTSG6; PRELIMINARY; PRI: 436 AA.

AC QRTSG6; 01-JUN-2002 (Trembl, 21, Created)

DI 01-JUN-2002 (Trembl, 21, Last sequence update)

DI 01-JUN-2002 (Trembl, 21, Last annotation update)

DE RKR-like protein.

GN RKR.

OS Biomphalaria glabrata (bloodfluke planorb).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;

OC Planorbidae; Biomphalaria.

OX NCBI_taxid=6526;

RN [1]

RP SEQUENCE FROM N.A.

RA Brown R., Mendoza R.L., Noel C., Escriba H., Cornette J., Landet V.,

RT "A conserved RKR family nuclear receptor from the mollusc Biomphalaria

glabrata.";

RI Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.

IR EMBL: AY048663; AAL86461.1; -

SV SEQUENCE 436 AA; 4825 MW; 94D6E84B5CEH4F9 CRC64;

Query Match 17.8%; Score 448.5; DB 5; Length 436;

Host Local Similarity 30.1%; Pred. No. 5,8e-33;

Matches 132; Conservative 63; Mismatches 170; Indels 7; Gaps 14;

```

QY      15 GPEPQTSNNVIMPTIPIHSPVVIHQQLSHIYAPVQKSPVCEAPSLHETLVNPEELKPK 71
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      55 GPGGMSMASSTQPSFG---POGMHSPGMHSPSTSMGSP-----PM----- 92
QY      75 VSGNRCASPVTSR-----GSKRAHFGAVNSVASYHYGVWSPGYKAFKPSIQG 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      93 ----LQLSP-SRPSRPNLPHSSITHKICAIQIPASGKHGVVSCDQKGFPRIVK 147
QY      127 HNIYTPALNCTIIKRNPKSYQVCPKGYTVVNVVCGSPRERGIVPLVPPGSADEQL 186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      148 DITVAIKRKNMTHQKPPQNPQYCPYMKQSMKMPKPAVQEPF-----QVFFKGDSEV 202
QY      187 WACVAPSSGAPVPELLPA-LPEQIVLTLEAPPHVLLSKVATFTTANIMRSL 215
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      203 E-----STGANNIMVDEQLLEALAVIKITTYITDQKDY-----TNI 242
QY      246 TKIAKELVHIMSKAKKIPGVHLSIDQVRLLESTVMEVIMMGLMRSIDHPKLLIAP 305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      243 GQADAKQIFIVMAKRIPIHTEILEPQVHLLFACWNELLTACEHSISMAKMDILLAT 302
QY      306 PVVLDPPGKGVGLLEIFDMLA--TISPEPELKIQHKRYGVKAMILLNSSMPVATVP 364
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      303 GIVHHSNAR-QAGVSTIFDPVLTILVAKMIPMKMKTETLSGLPAVLPNPAQCL-TVAV 360

```

```

QY      465 QIAHSSPKLAHLINAVIIALVWVIAKSGTSSQVQSNMPLANIIMLSVPHAS-NKGMHLL 424
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      361 QEVQDLREKVY-----ASLEEVTKS--RYPEERGFRAKILLKIPAKRSIGLKLEHLP 411
QY      425 NMCKNVVPYDILLEML 442
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      412 FFKLIDQDPTDIFLMBML 429

```

Search completed: June 16, 2003, 14:15:09
Job time : 32.487 secs

GenScore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using SW model

Run on: June 16, 2003, 04:20:57 : Search time 109.628 seconds

(without alignment)
505,642 Million total operations

Title: US-08-826-361A-6

Portlet score: 2203

Sequence: 1 MNSYSPNTNLEFGPPQPT QGSMELANLMIISVPEAP 416

Scoring table:

BLASTM62

Gapop 10 0, Gapext 0 5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:

908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Listing first 120 summaries

1: /S1US2/gqdata/geneseq/geneseq-emb1/AA1980.FAI.*
2: /S1US2/gqdata/geneseq/geneseq-emb1/AA1981.FAI.*
3: /S1US2/gqdata/geneseq/geneseq-emb1/AA1982.FAI.*
4: /S1US2/gqdata/geneseq/geneseq-emb1/AA1983.FAI.*
5: /S1US2/gqdata/geneseq/geneseq-emb1/AA1984.FAI.*
6: /S1US2/gqdata/geneseq/geneseq-emb1/AA1985.FAI.*
7: /S1US2/gqdata/geneseq/geneseq-emb1/AA1986.FAI.*
8: /S1US2/gqdata/geneseq/geneseq-emb1/AA1987.FAI.*
9: /S1US2/gqdata/geneseq/geneseq-emb1/AA1988.FAI.*
10: /S1US2/gqdata/geneseq/geneseq-emb1/AA1989.FAI.*
11: /S1US2/gqdata/geneseq/geneseq-emb1/AA1990.FAI.*
12: /S1US2/gqdata/geneseq/geneseq-emb1/AA1991.FAI.*
13: /S1US2/gqdata/geneseq/geneseq-emb1/AA1992.FAI.*
14: /S1US2/gqdata/geneseq/geneseq-emb1/AA1993.FAI.*
15: /S1US2/gqdata/geneseq/geneseq-emb1/AA1994.FAI.*
16: /S1US2/gqdata/geneseq/geneseq-emb1/AA1995.FAI.*
17: /S1US2/gqdata/geneseq/geneseq-emb1/AA1996.FAI.*
18: /S1US2/gqdata/geneseq/geneseq-emb1/AA1997.FAI.*
19: /S1US2/gqdata/geneseq/geneseq-emb1/AA1998.FAI.*
20: /S1US2/gqdata/geneseq/geneseq-emb1/AA1999.FAI.*
21: /S1US2/gqdata/geneseq/geneseq-emb1/AA2000.FAI.*
22: /S1US2/gqdata/geneseq/geneseq-emb1/AA2001.FAI.*
23: /S1US2/gqdata/geneseq/geneseq-emb1/AA2002.FAI.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	FR	ID	Description
1	2203	100 0	416	18	AAW43213
2	2203	100 0	416	23	AAW47837
3	2203	100 0	416	18	AAW43214
4	2203	100 0	416	23	AAW47836
5	2198	99 8	477	18	AAW43212
6	2198	99 8	477	23	AAW47834
7	2198	99 8	477	18	AAW47834
8	2198	99 8	477	23	AAW47834
9	2198	99 8	477	18	AAW43215
10	2198	99 8	477	23	AAW47838

11	2198	99.8	540	22	AAU27422	Human oestrogen re
12	2198	99.8	540	22	AAH60649	Human oestrogen re
13	2198	99.8	540	23	AAW47835	Full length estroa
14	2198	99.8	540	23	AAV07270	Human oestrogen re
15	2198	99.8	793	23	AAH80756	Amino acid sequen
16	2179	98.9	485	20	AAV01597	Human oestrogen re
17	1985	90.1	549	20	AAW94125	Mouse oestrogen re
18	1983	90.0	549	20	AAV04434	Human oestrogen re
19	1981	89.9	445	18	AAW14723	Human oestrogen re
20	1981	89.9	445	18	AAW14723	Human oestrogen re
21	1972.5	89.5	445	21	AAV01596	Human oestrogen re
22	1966	89.2	540	20	AAV04433	Human oestrogen re
23	1965	89.2	485	18	AAW14725	Human oestrogen re
24	1964	89.2	567	20	AAW98128	Mouse oestrogen re
25	1728.5	78.5	510	20	AAV04435	Human oestrogen re
26	1728.5	78.5	510	20	AAW98126	Mouse oestrogen re
27	1499	68.0	414	20	AAW98127	Human oestrogen re
28	1499	68.0	740	20	AAV04436	Human oestrogen re
29	1375.5	48.8	582	22	AAV78227	Human oestrogen re
30	1065	48.3	547	23	AAH81600	Human oestrogen re
31	1065	48.3	595	22	AAH84509	Human oestrogen re
32	1065	48.3	595	22	AAH84512	Human oestrogen re
33	1064	48.3	581	23	AAH83791	Human oestrogen re
34	1063	48.3	595	20	AAV21626	Human oestrogen re
35	1063	48.3	595	22	AAH84505	Human oestrogen re
36	1063	48.3	595	22	AAH84506	Human oestrogen re
37	1063	48.3	595	22	AAH84513	Human oestrogen re
38	1063	48.3	595	23	AAH76378	Human oestrogen re
39	1060	48.1	595	22	AAH84507	Human oestrogen re
40	1056.5	48.0	575	23	AAH19430	Human oestrogen re
41	1056	47.9	595	22	AAH84508	Human oestrogen re
42	1056	47.9	595	22	AAH84514	Human oestrogen re
43	1054	47.8	595	22	AAH84516	Human oestrogen re
44	1054	47.8	595	22	AAH84511	Human oestrogen re
45	1054	47.8	595	22	AAH84510	Human oestrogen re
46	1054	47.8	595	22	AAH84511	Human oestrogen re
47	1054	47.8	595	22	AAH84511	Human oestrogen re
48	1054	47.8	595	22	AAH84511	Human oestrogen re
49	1054	47.8	595	22	AAH84511	Human oestrogen re
50	1054	47.8	595	22	AAH84511	Human oestrogen re
51	1054	47.8	595	22	AAH84511	Human oestrogen re
52	1054	47.8	595	22	AAH84511	Human oestrogen re
53	1054	47.8	595	22	AAH84511	Human oestrogen re
54	1054	47.8	595	22	AAH84511	Human oestrogen re
55	1054	47.8	595	22	AAH84511	Human oestrogen re
56	1054	47.8	595	22	AAH84511	Human oestrogen re
57	1054	47.8	595	22	AAH84511	Human oestrogen re
58	1054	47.8	595	22	AAH84511	Human oestrogen re
59	1054	47.8	595	22	AAH84511	Human oestrogen re
60	1054	47.8	595	22	AAH84511	Human oestrogen re
61	1054	47.8	595	22	AAH84511	Human oestrogen re
62	1054	47.8	595	22	AAH84511	Human oestrogen re
63	1054	47.8	595	22	AAH84511	Human oestrogen re
64	1054	47.8	595	22	AAH84511	Human oestrogen re
65	1054	47.8	595	22	AAH84511	Human oestrogen re
66	1054	47.8	595	22	AAH84511	Human oestrogen re
67	1054	47.8	595	22	AAH84511	Human oestrogen re
68	1054	47.8	595	22	AAH84511	Human oestrogen re
69	1054	47.8	595	22	AAH84511	Human oestrogen re
70	1054	47.8	595	22	AAH84511	Human oestrogen re
71	1054	47.8	595	22	AAH84511	Human oestrogen re
72	1054	47.8	595	22	AAH84511	Human oestrogen re
73	1054	47.8	595	22	AAH84511	Human oestrogen re
74	1054	47.8	595	22	AAH84511	Human oestrogen re
75	1054	47.8	595	22	AAH84511	Human oestrogen re
76	1054	47.8	595	22	AAH84511	Human oestrogen re
77	1054	47.8	595	22	AAH84511	Human oestrogen re
78	1054	47.8	595	22	AAH84511	Human oestrogen re
79	1054	47.8	595	22	AAH84511	Human oestrogen re
80	1054	47.8	595	22	AAH84511	Human oestrogen re
81	1054	47.8	595	22	AAH84511	Human oestrogen re
82	1054	47.8	595	22	AAH84511	Human oestrogen re
83	1054	47.8	595	22	AAH84511	Human oestrogen re

XX 25-MAR-1997: 2001EP-0204021.
 XX
 XX 26-MAR-1996: 96EP-0200820.
 XX 22-NOV-1996: 96EP-0204284.
 XX 25-MAR-1997: 97EP-0200903.
 XX
 XX (AIKU) AKZO NOBEL NV.
 XX
 XX Mosselman S, Dijkema R:
 XX WPI: 2002-084444/12.
 XX N-PSDB: AAI72151.
 XX
 XX Now isolated chimeric receptor comprising a DNA binding domain and/or
 XX liquid binding domain of a new estrogen receptor, for identifying
 XX functional ligands or hormonal analogs for the receptor
 XX
 XX Example B: Page 17-18; 45pp; English.

XX The sequences given in AAB47836 and AAB47837 show splice variants
 XX of a novel estrogen receptor (ER). The gene encoding this new ER is
 XX located on chromosome 14 and has a different tissue distribution from
 XX classical ER. This ER also has two orphan ER's, ER-alpha and ER-beta.
 XX These orphan receptors have estrogen receptor related structure but do
 XX not appear to be able to bind estradiol or other ER ligands. These
 XX splice variants differ from the parent ER around exon 8. One clone
 XX contains exon 8b through alternative splicing. This causes an immediate
 XX termination of the reading thereby creating a truncation at the carboxy
 XX terminus. The other splice variant contains an alternative exon 8, exon
 XX 8c, which encodes two C-terminal amino acids and has a stop codon.
 XX The splice variant proteins do not contain the AF 2 region and therefore
 XX probably lack the ability to modulate transcription of target genes in
 XX a ligand dependant manner. However they may be able to interfere with
 XX the functioning of the WT classical ER or the novel ER of the invention,
 XX either by heterodimerisation or by occupying estrogen response elements
 XX or by interactions with other transcription factors.

XX Sequence 416 AA:

Query Match: 100.0%; Score 2203; DB 23; Length 416;
 Best Local Similarity 100.0%; Pred. No. 6; 3e-202;
 Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 YY 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 YY 121 KRSIQGHNHYTPATNCTTDRKPKRSYVACPLKRYVGVQVWVGSPPREPGYPLVPOQR 180
 121 KRSIQGHNHYTPATNCTTDRKPKRSYVACPLKRYVGVQVWVGSPPREPGYPLVPOQR 180
 121 KRSIQGHNHYTPATNCTTDRKPKRSYVACPLKRYVGVQVWVGSPPREPGYPLVPOQR 180
 121 KRSIQGHNHYTPATNCTTDRKPKRSYVACPLKRYVGVQVWVGSPPREPGYPLVPOQR 180
 YY 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 YY 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 YY 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416
 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416
 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416
 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416

RESULT 4
 AAB478214
 ID AAB478214 standard: protein; 418 AA.
 XX
 XX AAB478214:
 AC
 XX
 XX 20-APR-1998 (first entry)
 DI
 XX
 XX Human estrogen receptor protein splice variant (exon 8c).
 DE
 XX
 XX Estrogen receptor protein; steroid; alternative splicing; estradiol;
 KW
 XX
 XX estrogen; estril; screening.
 KM
 XX
 XX Homo sapiens.
 OS
 XX
 XX EP798378-A2.
 PN
 XX
 XX 01-OCT-1997.
 PD
 XX
 XX 25-MAR-1997: 97EP-0200903.
 PF
 XX
 XX 22-NOV-1996: 96EP-0204284.
 PR
 XX 26-MAR-1996: 96EP-0200820.
 PA
 XX (AIKU) AKZO NOBEL NV.
 XX
 XX Dijkema R, Mosselman S;
 XX WPI: 1997-473168/44.
 DE
 XX N-PSDB: AAI88414.
 XX
 XX DNA encoding estrogen receptor - useful in screening assay to
 FT
 XX identify novel ligands or hormonal analogues
 XX
 XX Claim 4: Page 30-32; 45pp; English.
 PS
 XX
 XX This sequence represents a splice variant of a novel estrogen binding
 CC protein isolated from human thymus tissue. This protein contains an
 CC alternative exon 8 (exon 8c) of the novel estrogen receptor represented
 CC in AAB478412. This novel receptor is able to bind and be activated by
 CC estradiol, estrone and estril, can be used in a screening assay for the
 CC identification of new drugs e.g. novel ligands or hormonal analogues.
 CC This variant does not contain an AF-2 region and therefore probably
 CC lacks the ability to modulate transcription of target genes in a ligand
 CC dependent fashion.
 CC
 XX
 XX Sequence 418 AA:

Query Match: 100.0%; Score 2203; DB 18; Length 418;
 Best Local Similarity 100.0%; Pred. No. 6; 3e-202;
 Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 1 MNSYSPSNVTNLEGGGROTTSPNVLMPFGHLSPLVYHROLSHLYAEPOKSPWCEAKSL 60
 YY 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 61 EHTLPVNRKELTKRVSQSNPQASVPTGNSKPDAPPAVAVSVDVASYHGVWSSCPGVAF 120
 YY 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 181 SAEQDHCACAKRSQSHAPVPELIDMSPEQVITLIEAPPHVILSPISAPPTAS 240
 YY 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 241 MMSSTKLAKKEIVHMSMAKKIPGVVLSIDQVLELSSQVMEVIMMIMWSIDHCK 300
 YY 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416
 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416
 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416
 361 VTAQDADSSRKLAHLINAVTALVAVTARSCISQSQSMPLANTMLISHPVAF 416

```

|||||
1b 11FAPPLVLDROEGKVEGILELLEMLATISPEHELMGHELEGVAMILLNSMYPL 600
yy 461 VIATGADSSSKLAHLNAVITAIYVIAKSSISGQOSMRANLMLLSHYRQR 416
1b 461 VIATGADSSSKLAHLNAVITAIYVIAKSSISGQOSMRANLMLLSHYRQR 416

RESULT 4
AA047846 standard: protein: 418 AA.
xx
xx AA047846:
xx
xx 25 MAR 2002 (first entry)
xx
xx ER splice variant ER alpha.
xx
xx DNA binding domain; DBD; ligand binding domain; LBD; chimera receptor;
xx estrogen receptor; ER; chromosome 14; ER-alpha; ER-beta; exon 8; PCR;
xx estradiol; nuclear receptor; progesterone receptor; amplify; primer;
xx polymerase chain reaction; AF-2 region; splice variant.
xx
xx Homo sapiens.
xx
xx E01162264 A2.
xx
xx 12 DEC 2001.
xx
xx 25 MAR 1997: 2001EP-0209021.
xx
xx 25 MAR 1996: 96EP-0209020.
xx
xx 22 NOV 1996: 96EP-0204284.
xx
xx 25 MAR 1997: 97EP-0209043.
xx
xx (AKU) AKZO NOBEL NV.
xx
xx Mosselman S; Djikema K;
xx
xx W01: 2002 084414/12.
xx
xx N PSDB: AA172146.
xx
xx New Isolated chimera receptor comprising a DNA binding domain and/or
xx ligand binding domain of a new estrogen receptor, for identifying
xx functional ligands or hormonal analogs for the receptor
xx
xx Example B: Page 24 25; 45pp; English.
xx
xx The sequences given in AA047846 and AA047847 show splice variants
xx of a novel estrogen receptor (ER). The gene encoding this new ER is
xx located on chromosome 14 and has a different tissue distribution from
xx classical ER. This ER also has two orphan ERs, ER alpha and ER-beta.
xx These orphan receptors have estrogen receptor related structure but do
xx not appear to be able to bind estradiol or other ER ligands. These
xx splice variants differ from the parent ER around exon 8. One clone
xx contains exon 8B through alternative splicing. This causes an immediate
xx termination of the reading thereby creating a truncation at the carboxy
xx terminus. The other splice variant contains an alternative exon 8, exon
xx 8C, which encodes two C-terminal amino acids and then has a stop codon.
xx The splice variant proteins do not contain the AF-2 region and therefore
xx probably lack the ability to modulate transcription of target genes in
xx a ligand dependent manner. However they may be able to interfere with
xx the functioning of the ER classical ER or the novel ER of the invention,
xx either by heterodimerisation or by occupying estrogen response elements
xx or by interactions with other transcription factors.
xx
xx Sequence 418 AA:
xx
xx Query Match: 100.0%; Score 2203; DB 24; Length 418.
xx Host Local Similarity 100.0%; Pred. No. 6, to 202;
xx Matches 416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
yy 1 MNSITSNVNLKRGKGTSTSNVIMPTWHLSPVYHQLSHLYAEVQKSPWCEARSL 60

```

```

|||||
1b 11FAPPLVLDROEGKVEGILELLEMLATISPEHELMGHELEGVAMILLNSMYPL 600
yy 61 EHLIVNRELKRVSNAGCASPVLGKSKRAHCVASVQASGYAVGWSLQCKRPF 120
1b 61 EHLIVNRELKRVSNAGCASPVLGKSKRAHCVASVQASGYAVGWSLQCKRPF 120
yy 121 KRSLQGHINLYGCVATNGTTLQKRRKRSVQATKLRKCYECVQKCSRRRQGYRIYRQR 180
1b 121 KRSLQGHINLYGCVATNGTTLQKRRKRSVQATKLRKCYECVQKCSRRRQGYRIYRQR 180
yy 181 SADEQLHFAKAKRSNAGHAPKVEELDALSPQSLVTLTLEAPPHVLSRPSAEPFVAS 240
1b 181 SADEQLHFAKAKRSNAGHAPKVEELDALSPQSLVTLTLEAPPHVLSRPSAEPFVAS 240
yy 241 MMSSTLKLAHKLAVIMLSMAKKLPGVELSLPDQVLESLWMVYLMGLMKRSIDHQR 400
1b 241 MMSSTLKLAHKLAVIMLSMAKKLPGVELSLPDQVLESLWMVYLMGLMKRSIDHQR 400
yy 301 LIFAPPLVLDROEGKVEGILELLEMLATISPEHELMGHELEGVAMILLNSMYPL 460
1b 301 LIFAPPLVLDROEGKVEGILELLEMLATISPEHELMGHELEGVAMILLNSMYPL 460
yy 461 VIATGADSSSKLAHLNAVITAIYVIAKSSISGQOSMRANLMLLSHYRQR 416
1b 461 VIATGADSSSKLAHLNAVITAIYVIAKSSISGQOSMRANLMLLSHYRQR 416

RESULT 5
AA047212 standard: protein: 477 AA.
xx
xx AA047212:
xx
xx 20 APR 1998 (first entry)
xx
xx Human estrogen receptor protein.
xx
xx estrogen receptor protein; steroid; alternative splicing; estradiol;
xx estrone; estradiol; screening.
xx
xx Homo sapiens.
xx
xx E0798578 A2.
xx
xx 01 OCT 1997.
xx
xx 25 MAR 1997: 97EP-0209043.
xx
xx 22 NOV 1996: 96EP-0204284.
xx
xx 25 MAR 1997: 96EP-0209043.
xx
xx (AKU) AKZO NOBEL NV.
xx
xx Djikema K; Mosselman S;
xx
xx W01: 1997 473168/44.
xx
xx N PSDB: AA188412.
xx
xx DNA encoding estrogen receptor useful in screening assay to
xx identify novel ligands or hormonal analogues
xx
xx Claim 4: Page 18-20; 45pp; English.
xx
xx This sequence represents a novel estrogen binding protein. The cDNA
xx sequence which encodes this protein can be alternatively spliced
xx resulting in the detection of additional transcripts (see AA188413).
xx This receptor is able to bind and be activated by estradiol, estrone and
xx estrone, can be used in a screening assay for the identification of new
xx drugs e.g. novel ligands or hormonal analogues.
xx
xx Sequence 477 AA:
xx

```

Query Match 99.8% Score 2198; DB 18; Length 477;
 Best Local Similarity 100.0%; Pred. No. 2, 4e-201;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MNYSTFNVNTNLEGGHGGTTSPTNVIWPTPSHISPLVHGOISHEIYAEKSPWTEARSL 60
 DB 1 MNYSTFNVNTNLEGGHGGTTSPTNVIWPTPSHISPLVHGOISHEIYAEKSPWTEARSL 60
 YY 61 EHLLPVNPEETIKPKVSNIPVAVSPVPSKPAHPTAVSYASGVYGVWVSCPGCAAF 120
 DB 61 EHLLPVNPEETIKPKVSNIPVAVSPVPSKPAHPTAVSYASGVYGVWVSCPGCAAF 120
 YY 121 KRSTIGCHNHYTPATWCTTOKNPKSCGACPLPKCYGVGWKCSPEPCGYPIVPRGR 180
 DB 121 KRSTIGCHNHYTPATWCTTOKNPKSCGACPLPKCYGVGWKCSPEPCGYPIVPRGR 180
 YY 181 SAEPQIHCAKAKRPSSTHAPVPEIITPAISPEQVITILEAPPHVILSPHSAPFTFAS 240
 DB 181 SAEPQIHCAKAKRPSSTHAPVPEIITPAISPEQVITILEAPPHVILSPHSAPFTFAS 240
 YY 241 MMSSTLRKADKELVHMSWAKKIPGVFELSLPDVYALLFSCWMEVIMMGLMWRSTIDHCK 300
 DB 241 MMSSTLRKADKELVHMSWAKKIPGVFELSLPDVYALLFSCWMEVIMMGLMWRSTIDHCK 300
 YY 301 LIFAPDLVIDRODCKVEGILETFPMGLATTSRPFELKIQHKEYICVAKAMILLNSMYPL 360
 DB 301 LIFAPDLVIDRODCKVEGILETFPMGLATTSRPFELKIQHKEYICVAKAMILLNSMYPL 360
 YY 361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMRILMLLSHYRHA 415
 DB 361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMRILMLLSHYRHA 415

RESULT 6
 AAB47834
 10 AAB47834 standard; Protein: 477 AA.
 AC AAB47834:
 01 25-MAR-2002 (first entry)
 XX
 DE Estrogen receptor.
 XX
 KW DNA binding domain; DBD; ligand binding domain; LBD; chimeric receptor;
 KW estrogen receptor; ER; chromosome 14; ER-alpha; ER-beta; exon 8;
 KW estradiol; nuclear receptor; progesterone receptor.
 XX
 OS Homo sapiens.
 EN EP1162264-A2
 XX
 ID 12-DEC-2001.
 XX
 PF 25-MAR-1997; 2001EP-0202021.
 XX
 PR 25-MAR-1996; 96EP-0200820.
 PR 22-NOV-1996; 96EP-020284.
 PR 25-MAR-1997; 97EP-0200903.
 XX
 PA (AIKU) AKZO NOBEL, NV.
 XX
 PI Mosselman S, Di Kema R;
 XX
 DE WP1: 2003-084414/12.
 DE N-PSDB: AA172128.
 XX
 PT New isolated chimeric receptor comprising a DNA binding domain and/or
 PT ligand binding domain of a new estrogen receptor, for identifying
 PT functional ligands or hormonal analogs for the receptor.
 XX
 PS Example A; Page 15-17; 35pp; English.
 CC This sequence shows a novel estrogen receptor (ER). The gene encoding

CC This new ER is located on chromosome 14 and has a different tissue
 CC distribution from classical ER. This ER also has two orphan ER's,
 CC ER alpha and ER beta. These orphan receptors have estrogen receptor
 CC related structure but do not appear to be able to bind estradiol or other
 CC ER ligands. The DNA binding domain (DBD) and ligand binding domain (LBD)
 CC from this ER may be used in the chimeric receptor of the invention
 CC which also has an N-terminal domain. The chimeric receptor, or DNA
 CC encoding it, is useful in a screening assay for identification of new
 CC drugs. Similar chimeric receptors comprising the LBD of the new ER,
 CC and also comprising the DBD and an N-terminal domain derived from
 CC another nuclear receptor e.g., progesterone receptor, are useful for the
 CC screening of compounds to identify new ligands or hormone analogs which
 CC are able to activate the new ER. Chimeric receptors comprising a DBD
 CC of the new ER, and LBD and an N-terminal domain derived from
 CC another nuclear receptor, can be used to identify new ligands or
 CC hormone analogs for the nuclear receptors.

Sequence 477 AA:

Query Match 99.8% Score 2198; DB 24; Length 477;
 Best Local Similarity 100.0%; Pred. No. 2, 4e-201;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 1 MNYSTFNVNTNLEGGHGGTTSPTNVIWPTPSHISPLVHGOISHEIYAEKSPWTEARSL 60
 DB 1 MNYSTFNVNTNLEGGHGGTTSPTNVIWPTPSHISPLVHGOISHEIYAEKSPWTEARSL 60
 YY 61 EHLLPVNPEETIKPKVSNIPVAVSPVPSKPAHPTAVSYASGVYGVWVSCPGCAAF 120
 DB 61 EHLLPVNPEETIKPKVSNIPVAVSPVPSKPAHPTAVSYASGVYGVWVSCPGCAAF 120
 YY 121 KRSTIGCHNHYTPATWCTTOKNPKSCGACPLPKCYGVGWKCSPEPCGYPIVPRGR 180
 DB 121 KRSTIGCHNHYTPATWCTTOKNPKSCGACPLPKCYGVGWKCSPEPCGYPIVPRGR 180
 YY 181 SAEPQIHCAKAKRPSSTHAPVPEIITPAISPEQVITILEAPPHVILSPHSAPFTFAS 240
 DB 181 SAEPQIHCAKAKRPSSTHAPVPEIITPAISPEQVITILEAPPHVILSPHSAPFTFAS 240
 YY 241 MMSSTLRKADKELVHMSWAKKIPGVFELSLPDVYALLFSCWMEVIMMGLMWRSTIDHCK 300
 DB 241 MMSSTLRKADKELVHMSWAKKIPGVFELSLPDVYALLFSCWMEVIMMGLMWRSTIDHCK 300
 YY 301 LIFAPDLVIDRODCKVEGILETFPMGLATTSRPFELKIQHKEYICVAKAMILLNSMYPL 360
 DB 301 LIFAPDLVIDRODCKVEGILETFPMGLATTSRPFELKIQHKEYICVAKAMILLNSMYPL 360
 YY 361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMRILMLLSHYRHA 415
 DB 361 VTATQADSSRKLAHLINAVTDALVWVIAKSGISSQOOSMRILMLLSHYRHA 415

RESULT 7
 AAM14724
 10 AAM14724 standard; Protein: 485 AA.
 AC AAM14724:
 01 08-JUN-1997 (first entry)
 XX
 DE Human oestrogen receptor beta (ER-beta).
 XX
 KW cytochrome P-450; estrogen receptor beta; ER-beta;
 KW nuclear receptor; prostate cancer; breast prostatic hyperplasia;
 KW ovarian cancer; cardiovascular disease; osteoporosis;
 KW environment; pollutant.
 XX
 OS Homo sapiens.
 EN
 XX
 DE Key
 DE Domain
 FT Location/Qualifiers
 FT 104..169
 FT /label=DNA-binding_domain
 FT 260..457

FT /Label: ligand binding domain
 XX W070948 A2.
 FN 15 MAR 1997
 XX
 XX
 XX 09 SEP 1996: 96W0-EPO4943.
 XX
 XX 08 MAY 1996: 96G28-0009576.
 XX 08 SEP 1995: 95G38-0014672.
 XX 15 MAR 1996: 96G38-0005550.
 XX 11 APR 1996: 96G38-0007532.
 XX (KAGO) KAGO BIO AB.
 XX
 XX Kenneth E. Gustafsson J. Kuiper GC:
 XX W071997 192842/17.
 XX N P508: AAT62843.
 XX
 XX New isolated oestrogen receptor beta, used to develop prods. for
 FT treating e.g. cancers, CNS diseases, osteoporosis or cardiovascular
 FT disease.
 XX
 XX claim 1: Fig 1A: 45pp: English.
 XX
 XX A novel human orphan receptor (AAW14724) is related to the known
 XX oestrogen receptor ER alpha, and has been designated ER beta.
 XX It is an oestrogen receptor related nuclear receptor. The ER-beta
 XX amino acid sequence was deduced from a cDNA clone (AA162843) isolated
 XX from a human testis cDNA library. Rat, human and mouse ER-beta
 XX (AAW14724-25) can be used to isolate molecules for use in the
 XX treatment of cardiovascular diseases, central nervous system
 XX diseases, osteoporosis, prostate or ovarian cancer or benign
 XX prostate hyperplasia and to test environmental chemicals for
 XX oestrogenic activity.
 XX
 XX Sequence: 485 AA:
 Query Match 99.88% Score 2198: DB 18: Length 485:
 Best Local Similarity 100.0%: Prod. No. 2-36-203:
 Matches 415: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 YY 1 MNSSTISNVTNLEGGGRQTTSTNVTNPTPGHISPLVYHROLSHLIYAEPOKSPWEASTL 60
 DB 9 MNSSTISNVTNLEGGGRQTTSTNVTNPTPGHISPLVYHROLSHLIYAEPOKSPWEASTL 68
 YY 61 EHTLPVNRKELTKKRVASNGCASVYTGSGKRDAHFCVAGSYDASGYHGWSCGCKKAF 120
 DB 69 EHTLPVNRKELTKKRVASNGCASVYTGSGKRDAHFCVAGSYDASGYHGWSCGCKKAF 128
 YY 121 KRSLGCHNDYICVATNCTIDKRRKRSYVAELKRCYEWGMAKSGRRRCGYRIVRQR 180
 DB 129 KRSLGCHNDYICVATNCTIDKRRKRSYVAELKRCYEWGMAKSGRRRCGYRIVRQR 188
 YY 181 SADEOLHCAKAKRSQGHAPRVRELLDALSPBOVLTLIAEPHHVLSRISAPFTVAS 240
 DB 189 SADEOLHCAKAKRSQGHAPRVRELLDALSPBOVLTLIAEPHHVLSRISAPFTVAS 248
 YY 241 MMSSTLTKLAKELVHMI SWAKKIPGVVELSTFDYKRLLESYWMVEYLMGCLMHRSLDIDRK 400
 DB 249 MMSSTLTKLAKELVHMI SWAKKIPGVVELSTFDYKRLLESYWMVEYLMGCLMHRSLDIDRK 408
 YY 401 LIFAPDLVLDROGKCVESLLEFDMLATTSRPFELKLGKREYLCYKAMILLNSMYPL 460
 DB 409 LIFAPDLVLDROGKCVESLLEFDMLATTSRPFELKLGKREYLCYKAMILLNSMYPL 468
 YY 461 VTATGVDADSSRKLAHLINAVTDALVWVIAKSGTSSQGSQSMRIANLMLLSHVRIA 415
 DB 469 VTATGVDADSSRKLAHLINAVTDALVWVIAKSGTSSQGSQSMRIANLMLLSHVRIA 423

AAW44215
 XX 10 AAW44215 standard: Protein: 530 AA.
 XX
 XX AAW44215:
 XX 20-APR-1998 (first entry)
 XX
 XX human oestrogen receptor protein #2.
 DE
 XX oestrogen receptor protein: steroid alternative splicing: estradiol:
 KW oestrogen estradiol: screening.
 XX
 XX Homo sapiens.
 XX
 XX E0798478-A2.
 XX
 XX 01-OCT-1997.
 XX
 XX 25-MAR-1997: 97EP-0209903.
 XX
 XX 22 NOV 1996: 96EP-0204284.
 XX 26 MAR 1996: 96EP-0208820.
 XX
 XX (AIKU) AKZO NOBEL NV.
 XX
 XX Dijkema R. Mosselman S:
 XX
 XX W071997 473188/44.
 XX N P508: AAT68415.
 XX
 XX This sequence represents a novel oestrogen binding protein isolated from
 XX human testis cDNA in order to study upstream translation initiation
 XX exons using 5' RACE-pick technology. This receptor is able to bind and
 XX be activated by estradiol, estrone and estradiol, can be used in a screening
 XX assay for the identification of new drugs e.g. novel ligands or hormonal
 XX analogues.
 XX
 XX Sequence: 530 AA:
 Query Match 99.88% Score 2198: DB 18: Length 530:
 Best Local Similarity 100.0%: Prod. No. 2-76-203:
 Matches 415: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 YY 1 MNSSTISNVTNLEGGGRQTTSTNVTNPTPGHISPLVYHROLSHLIYAEPOKSPWEASTL 60
 DB 54 MNSSTISNVTNLEGGGRQTTSTNVTNPTPGHISPLVYHROLSHLIYAEPOKSPWEASTL 113
 YY 61 EHTLPVNRKELTKKRVASNGCASVYTGSGKRDAHFCVAGSYDASGYHGWSCGCKKAF 120
 DB 114 EHTLPVNRKELTKKRVASNGCASVYTGSGKRDAHFCVAGSYDASGYHGWSCGCKKAF 173
 YY 121 KRSLGCHNDYICVATNCTIDKRRKRSYVAELKRCYEWGMAKSGRRRCGYRIVRQR 180
 DB 174 KRSLGCHNDYICVATNCTIDKRRKRSYVAELKRCYEWGMAKSGRRRCGYRIVRQR 233
 YY 181 SADEOLHCAKAKRSQGHAPRVRELLDALSPBOVLTLIAEPHHVLSRISAPFTVAS 240
 DB 234 SADEOLHCAKAKRSQGHAPRVRELLDALSPBOVLTLIAEPHHVLSRISAPFTVAS 293
 YY 241 MMSSTLTKLAKELVHMI SWAKKIPGVVELSTFDYKRLLESYWMVEYLMGCLMHRSLDIDRK 400
 DB 294 MMSSTLTKLAKELVHMI SWAKKIPGVVELSTFDYKRLLESYWMVEYLMGCLMHRSLDIDRK 453
 YY 401 LIFAPDLVLDROGKCVESLLEFDMLATTSRPFELKLGKREYLCYKAMILLNSMYPL 460
 DB 454 LIFAPDLVLDROGKCVESLLEFDMLATTSRPFELKLGKREYLCYKAMILLNSMYPL 413
 YY 461 VTATGVDADSSRKLAHLINAVTDALVWVIAKSGTSSQGSQSMRIANLMLLSHVRIA 415

```

ld 414 VTATGDAISSPKAHINAVTIAVWIAVAKSSISSQGSMPANIMTISHVPHA 148
|||||

RESULT 9
AAW7858 standard; protein; 530 AA.
XX AAW7858:
AC AAW7858:
DT 07-JUN-1999 (first entry)
XX
XX Human oestrogen receptor-beta.
XX
XX oestrogen receptor-beta; hkr-beta; human.
XX
XX Homo sapiens.
XX
XX W09907847-A1.
XX
XX 18-FEB-1999.
XX
XX 20-JUL-1998; 98W-0514944.
XX
XX 05-AUG-1997; 97US-0906165.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX What RA, Henderson RA, Hsiao C, Karathanasis SK;
XX
XX WPI: 1999-167424/14.
XX
XX N-PSDB: AAX24164.
XX
XX
XX Human oestrogen receptor beta coding sequence - useful in the
XX production of human oestrogen receptor-beta and identification of
XX human oestrogen receptor-beta interactive compounds
XX
XX Claim 17; Page 42-44; 49pp; English.
XX
XX This polypeptide comprises a full-length human oestrogen receptor beta
XX (hkr-beta), including 45 previously unknown N-terminal amino acid
XX residues that are believed to contribute to the transcription
XX activation function of the receptor. The amino acid sequence was
XX deduced from the nucleotide sequence (see AAX24164) of an isolated
XX cDNA clone. hkr-beta is selectively expressed in the thymus,
XX spleen, ovary and testes. The invention encompasses hkr-beta
XX polypeptides and polypeptides, particularly peptides which
XX include residues 1-45 of hkr-beta. The invention also provides
XX expression systems in which transcriptionally active hkr-beta or
XX fragments can be produced, as well as screening methods for
XX identifying hkr-beta agonists and antagonists (including
XX tissue-specific oestrogens and anti-oestrogens), and hkr-beta
XX co-activators and inhibitors.
XX
XX Sequence 530 AA.
XX
XX Query Match: 99.8%, Score 2198, DB 20, Length 530,
XX Best Local Similarity 100.0%; Prod. No. 2-76-201;
XX Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

14 244 SAKPGTFRWAGKAPKSGAPVPPILALALSPQGLVETIPAPPHVITSPGAPETAS 294
|||||
15 241 MMMSLTKLAKKEIVHMI SWAKKTPPEVELSPTVPLLESCOMMEVIMMGLMWSCHPPCK 300
|||||
16 294 MMMSLTKLAKKEIVHMI SWAKKTPPEVELSPTVPLLESCOMMEVIMMGLMWSCHPPCK 354
|||||
17 301 LIFAPDLVLDKKEGKVEGLLELHMLALISPEELNGLKELYVYAMILLNSMYPL 360
|||||
18 354 LIFAPDLVLDKKEGKVEGLLELHMLALISPEELNGLKELYVYAMILLNSMYPL 414
|||||
19 461 VTATGDAISSPKAHINAVTIAVWIAVAKSSISSQGSMPANIMTISHVPHA 415
|||||
20 414 VTATGDAISSPKAHINAVTIAVWIAVAKSSISSQGSMPANIMTISHVPHA 468
|||||

```

```

RESULT 10
AAE10457
ID AAE10457 standard; protein; 530 AA.
XX
XX AAE10457:
XX
XX 18-DEC-2001 (first entry)
XX
XX
XX Human estrogen receptor (ER)-beta protein.
XX
XX Estrogen receptor-beta; dietary compound; environmental pollutant;
XX red wine; ER-beta mediated cell growth inhibition modulator; human.
XX
XX Homo sapiens.
XX
XX W0200169262-A1.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001W-US08276.
XX
XX 15-MAR-2000; 2000US-189605P.
XX
XX (UYMA-) UNIV MASSACHUSETTS.
XX
XX Ho S;
XX
XX WPI: 2001-596920/67.
XX
XX N-PSDB: AAD18442.
XX
XX
XX In vitro screening for modulator of estrogen receptor-beta-mediated
XX cell growth inhibition by contacting cell having functional ER-beta
XX protein with candidate compound and detecting ER-beta regulated gene
XX expression.
XX
XX Example 3; Fig 7A; 49pp; English.
XX
XX The invention relates to in vitro screening for modulator of estrogen
XX receptor-beta (ER-beta)-mediated cell growth inhibition. The method
XX involves contacting mammalian cell having functional ER-beta protein
XX with candidate compound and detecting increase/decrease of ER-beta-
XX regulated gene expression in presence of candidate compound compared
XX to expression in absence of the compound. The method is useful for
XX identifying a compound that modulates ER-beta mediated cell growth
XX inhibition. It is useful for testing dietary compounds, e.g. red wine,
XX for the presence of estrogen mimics that trigger ER-beta-mediated
XX processes, thus affecting physiological function positively or
XX negatively and for testing environmental pollutants for the presence
XX of estrogen mimics that may pose health risks involving ER-beta
XX mediated processes. The present sequence is human ER-beta protein.
XX
XX Sequence 530 AA:
XX
XX Query Match: 99.8%, Score 2198, DB 22, Length 530;
XX Best Local Similarity 100.0%; Prod. No. 2-76-201;
XX Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


CC intensity in the measured fluorescence. The analyzers are for studying
CC behaviors of fluorescence-labeled molecules particularly intracellular
CC biological molecules like proteins at molecular level, e.g. proteins
CC functions and interactions. Such analyzers are stable and convenient to
CC operate, thereby enabling easy performance of fluorescence correlation
CC spectroscopy, fluorescence intensity distribution analysis and
CC fluorescence intensity multiple distribution analysis. The present
CC sequence represents the amino acid sequence of a fusion protein
CC comprising a GFP (green fluorescent protein) and hER (human estrogen
CC receptor) beta protein.

XX Sequence 793 AA:

Query Match 99.8%; Score 2158; DB 23; Length 793;

Best Local Similarity 100.0%; Prod. No. 4.7e-201;

Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

UY 1 MNSYSPSNVTNLEGGPGPOTTSNNVIMPTPSHISPLVNHQISHLVAPRQSPWFAEKL 60
    |||
DB 117 MNSYSPSNVTNLEGGPGPOTTSNNVIMPTPSHISPLVNHQISHLVAPRQSPWFAEKL 376
    |||
UY 61 FHTLVNRPETLKKKVSNGPASPVTGPGSPALFVAGSYAGYAGVWSPGZGAF 120
    |||
DB 377 FHTLVNRPETLKKKVSNGPASPVTGPGSPALFVAGSYAGYAGVWSPGZGAF 436
    |||
UY 121 KRSLGHNHYTPPATNGCTTHKNNRKSQAGCPKPGYFVGVKSGSPFGGYPVLPQR 180
    |||
DB 437 KRSLGHNHYTPPATNGCTTHKNNRKSQAGCPKPGYFVGVKSGSPFGGYPVLPQR 496
    |||
UY 181 SADEQLHCAKAKRSGSHAPRVRELLDALSPQVLTLLEAPRHVLSRPSAPFTAS 240
    |||
DB 497 SADEQLHCAKAKRSGSHAPRVRELLDALSPQVLTLLEAPRHVLSRPSAPFTAS 556
    |||
UY 241 MMSLTKLADKELVHMTSMARKPGVEVLSFTQVRIESTVMEVLMKLMKRSIDHFGK 300
    |||
DB 557 MMSLTKLADKELVHMTSMARKPGVEVLSFTQVRIESTVMEVLMKLMKRSIDHFGK 615
    |||
UY 301 LIFAPDLVLDREGKVEGILEIFDMLATTSRPEELKLGKEYLCVAMILLNSMYPL 360
    |||
DB 617 LIFAPDLVLDREGKVEGILEIFDMLATTSRPEELKLGKEYLCVAMILLNSMYPL 676
    |||
UY 361 VTATGADSSPKLAHLNAVTALVWIAKSGISSQOOSMRLANIMLSHVHRA 415
    |||
DB 677 VTATGADSSPKLAHLNAVTALVWIAKSGISSQOOSMRLANIMLSHVHRA 731
    |||

```

RESULT 16

AAV01597 standard: Protein; 485 AA.

XX AAV01597:

DT 17-JUN-1999 (first entry)

DB An estrogen receptor-beta protein.

XX differential estrogen receptor: ER-alpha; ER-beta; activation;

KW APl sites: therapy.

XX Homo sapiens.

XX W09911760-A1.

XX 11-MAR-1999.

XX 31-AUG-1998: 98W0-US18030.

XX 04-SEP-1997: 97US-0923708

XX (RESO) UNIV CALIFORNIA.

XX Gustafsson J, Kuiper GJM, Kushner PJ, Nilsson S;

PI Paech K, Scanlan TS, Webb P;

XX W01: 1999 205173/17.

XX N-PSDB: AAX26866.

PT Screening for test compounds which inhibit or activate estrogen

PS receptor beta (ER beta) mediated activation at a APl site - for use

XX in therapy.

XX Claim 14: Page 43-44: 61pp; English.

XX The present sequence represents an estrogen receptor-beta. The

XX specification describes a method for screening test compounds for

XX differential estrogen receptor (ER) alpha-mediated and ER beta-mediated

XX activation at an APl site. The method is used for identifying compounds

XX which can be used in therapy.

XX Sequence 485 AA:

Query Match 98.9%; Score 2179; DB 20; Length 485;

Best Local Similarity 99.3%; Prod. No. 1.5e-199;

Matches 412; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

UY 1 MNSYSPSNVTNLEGGPGPOTTSNNVIMPTPSHISPLVNHQISHLVAPRQSPWFAEKL 60
    |||
DB 9 MNSYSPSNVTNLEGGPGPOTTSNNVIMPTPSHISPLVNHQISHLVAPRQSPWFAEKL 68
    |||
UY 61 FHTLVNRPETLKKKVSNGPASPVTGPGSPALFVAGSYAGYAGVWSPGZGAF 120
    |||
DB 9 GHTLVNRPETLKKKVSNGPASPVTGPGSPALFVAGSYAGYAGVWSPGZGAF 128
    |||
UY 121 KRSLGHNHYTPPATNGCTTHKNNRKSQAGCPKPGYFVGVKSGSPFGGYPVLPQR 180
    |||
DB 129 KRSLGHNHYTPPATNGCTTHKNNRKSQAGCPKPGYFVGVKSGSPFGGYPVLPQR 188
    |||
UY 181 SADEQLHCAKAKRSGSHAPRVRELLDALSPQVLTLLEAPRHVLSRPSAPFTAS 240
    |||
DB 189 SADEQLHCAKAKRSGSHAPRVRELLDALSPQVLTLLEAPRHVLSRPSAPFTAS 248
    |||
UY 241 MMSLTKLADKELVHMTSMARKPGVEVLSFTQVRIESTVMEVLMKLMKRSIDHFGK 300
    |||
DB 249 MMSLTKLADKELVHMTSMARKPGVEVLSFTQVRIESTVMEVLMKLMKRSIDHFGK 308
    |||
UY 301 LIFAPDLVLDREGKVEGILEIFDMLATTSRPEELKLGKEYLCVAMILLNSMYPL 360
    |||
DB 309 LIFAPDLVLDREGKVEGILEIFDMLATTSRPEELKLGKEYLCVAMILLNSMYPL 368
    |||
UY 361 VTATGADSSPKLAHLNAVTALVWIAKSGISSQOOSMRLANIMLSHVHRA 415
    |||
DB 369 VTATGADSSPKLAHLNAVTALVWIAKSGISSQOOSMRLANIMLSHVHRA 423
    |||

```

RESULT 17

AAW98125 standard: Protein; 549 AA.

XX AAW98125:

DT 21-JUN-1999 (first entry)

DB Mouse estrogen receptor beta isoform MEK-beta-1.

XX differential estrogen receptor: ER-beta-1; ER-beta-2; mouse; tumour;

KW diagnosis: therapy; isoform.

XX Mus musculus.

XX W09905170-A1.

XX 04-FEB-1999.

XX 28-JUL-1998: 98W0-US15639.

XX 30-JUL-1997: 97US-0054210.

24 JUN 1997: 970S 005486.
 XX
 PA (UYVA) UNIV YALE.
 XX
 E1 Denton R, Fraston C, Dalton R, Kisch-Kirou M.
 XX
 M01: 1999 142852/12.
 XX
 P1 Human oestrogen receptor beta 4 useful in treating and monitoring
 P2 tumours.
 PS
 XX Example 2: Fig 2a: 89pp: English.
 XX
 CC The present sequence is alternately spliced isoform mkr beta-1
 CC of the murine complete oestrogen receptor beta (ER beta-c), or
 CC mkr beta-1). The mkr beta-1 gene (see AAX249435) contains 9 exons,
 CC including the newly described exon 5b, which is not found in ER
 CC beta-1). Exon 5b is located in the ligand binding domain of mkr
 CC beta-1, and probably plays a significant role in mkr beta-1 function.
 CC The mkr beta-1 clone was isolated by RT-PCR of mRNA, and by
 CC amplification from a mouse embryonic stem cell genomic DNA library.
 CC It lacks exons 5b. Preliminary data indicate that the mkr beta-1
 CC isoform may be more active than the full-length mkr beta-1 (see
 CC AAW98128). 2 other isoforms, mkr beta-2 (see AAW98126) and rat ER
 CC beta-4 (see AAW98127), have also been isolated, and the human exon 5b
 CC region (see AAX24944) has been identified. Based on these sequences,
 CC the invention provides methods for identifying agents that block or
 CC augment ER beta c mediated transcriptional regulation, methods to
 CC determine whether ER beta c or its isoforms is being expressed in
 CC tissues or cells, and methods of identifying and using agents that
 CC block the transcriptional regulation of genes by ER beta c or its
 CC isoforms, which in turn modulates other biological and pathological
 CC processes. Gene therapy expression of ER beta-c is envisaged.
 XX
 SQ Sequence 549 AA;
 CC
 CC Query Match 90.1%; Score 1985; 108 20; Length 549;
 CC Best Local Similarity 89.4%; Pred. No. 6, 66 181;
 CC Matches 470; Conservative 20; Mismatches 24; Indels 0; Gaps 0.
 CC
 CC 1 MNSSTSNVNLGGTGTSTSNVLMPTSHLVVHQSLSHVYAEKSPWFEARSL 60
 CC 73 MNSVSVSTGTEEGVGTASINVLMTPSLATHGUSLGYAEKSPWFEARSL 142
 CC
 CC 61 EHTLVNRRGTLKRVSGNACASVIGTSSKRIAHCAVSSQASGYRQWSTGCKAPF 120
 CC 143 EHTLVNRRGTLKRVSGNACASVIGTSSKRIAHCAVSSQASGYRQWSTGCKAPF 192
 CC
 CC 121 KRSIGHNINYTGATNGCTDKNRRKSTGACRLKCYGVGWKGSRRRGTGYRVRQR 180
 CC 193 KRSIGHNINYTGATNGCTDKNRRKSTGACRLKCYGVGWKGSRRRGTGYRVRQR 252
 CC
 CC 181 SAEQDJDCAKAKRSAGAAKVEELDLALSPQIVLTLEAEPPVILSRISAPETPAS 240
 CC 253 SASEGVHCLNKAKRTSGHTPRKVELLNSLSPQIVLTLEAEPPVILSRISAPETPAS 312
 CC
 CC 241 MMSSTLKLAKRELVIHMTSWAKKIPGVVELSLDGVRLLESWMVEYLMGLMKRSIDHCK 400
 CC 413 MMSSTLKLAKRELVIHMTSWAKKIPGVVELSLDGVRLLESWMVEYLMGLMKRSIDHCK 472
 CC
 CC 401 LIFAPDLVLDROGKVEGILEPMLLATTSSREELKQREYLCYKAMILLNSMYPL 460
 CC 473 LIFAPDLVLDROGKVEGILEPMLLATTARPEELKQREYLCYKAMILLNSMYPL 432
 CC
 CC 461 VITATDADSSKRLAHLLNAVITATVAVLAKSSISQOOSMRANLMLLSHVRH 414
 CC 433 ATVSQFAESSRKLTHTLNAVITATVAVLAKSSISQOOSVRLANLMLLSHVRH 486
 CC
 CC RESULT 18
 CC AAY044
 CC ID AAY044 64 Standard: Protobin: 549 AA.
 CC XX

AAY04434;
 XX
 CC 25 JUN 1999 (first entry)
 XX
 DE Murine mkr beta-1 clone protein sequence.
 XX
 KW Estrogen receptor beta: oestrogen receptor beta; ER beta-
 XX
 OS Mus Sp.
 XX
 P1 W9905171 AL.
 P2 04-FEB-1999.
 P3 28-JUN-1998: 98W0 US15540.
 P4 40-JUN-1997: 970S-0054210.
 P5 28-JUN-1997: 970S-0054869.
 PA (UYVA) UNIV YALE.
 P1 Barton R, Chambon P, Denton R, Dupont S, Garnier J.
 P2 M01: 1999 142853/12.
 DE
 CC Novel murine oestrogen receptor-beta genes and splice variants -
 CC useful for treating oestrogen receptor-beta mediated disorders
 CC claim 9; Fig 2a: 89pp: English.
 CC
 CC The present invention describes the murine complete oestrogen receptor
 CC beta (mkr beta-c) also designated mkr beta-1). The present sequence
 CC represents mkr beta-1, which is an isoform of mkr beta-c. The protein
 CC can be used to screen for agents that modulate or block the interaction
 CC between the ER beta c protein and physiological, stilbene or
 CC triphenylethylene (anti)oestrogens. The antibody which binds the protein
 CC can be used for assays to detect activation of the ER beta protein and
 CC can be used for subcellular localisation of the protein. The
 CC complementary nucleic acid can be used to inhibit expression of the
 CC ER beta c nucleic acid.
 XX
 SQ Sequence 549 AA;
 CC
 CC Query Match 90.0%; Score 1984; 108 20; Length 549;
 CC Best Local Similarity 89.1%; Pred. No. 16 180;
 CC Matches 469; Conservative 21; Mismatches 24; Indels 0; Gaps 0;
 CC
 CC 1 MNSSTSNVNLGGTGTSTSNVLMPTSHLVVHQSLSHVYAEKSPWFEARSL 60
 CC 73 MNSVSVSTGTEEGVGTASINVLMTPSLATHGUSLGYAEKSPWFEARSL 142
 CC
 CC 61 EHTLVNRRGTLKRVSGNACASVIGTSSKRIAHCAVSSQASGYRQWSTGCKAPF 120
 CC 143 EHTLVNRRGTLKRVSGNACASVIGTSSKRIAHCAVSSQASGYRQWSTGCKAPF 192
 CC
 CC 121 KRSIGHNINYTGATNGCTDKNRRKSTGACRLKCYGVGWKGSRRRGTGYRVRQR 180
 CC 193 KRSIGHNINYTGATNGCTDKNRRKSTGACRLKCYGVGWKGSRRRGTGYRVRQR 252
 CC
 CC 181 SAEQDJDCAKAKRSAGAAKVEELDLALSPQIVLTLEAEPPVILSRISAPETPAS 240
 CC 253 SASEGVHCLNKAKRTSGHTPRKVELLNSLSPQIVLTLEAEPPVILSRISAPETPAS 312
 CC
 CC 241 MMSSTLKLAKRELVIHMTSWAKKIPGVVELSLDGVRLLESWMVEYLMGLMKRSIDHCK 400
 CC 413 MMSSTLKLAKRELVIHMTSWAKKIPGVVELSLDGVRLLESWMVEYLMGLMKRSIDHCK 472
 CC
 CC 401 LIFAPDLVLDROGKVEGILEPMLLATTSSREELKQREYLCYKAMILLNSMYPL 460
 CC 473 LIFAPDLVLDROGKVEGILEPMLLATTARPEELKQREYLCYKAMILLNSMYPL 432
 CC
 CC 461 VITATDADSSKRLAHLLNAVITATVAVLAKSSISQOOSMRANLMLLSHVRH 414
 CC 433 ATVSQFAESSRKLTHTLNAVITATVAVLAKSSISQOOSVRLANLMLLSHVRH 486
 CC
 CC RESULT 18
 CC AAY044
 CC ID AAY044 64 Standard: Protobin: 549 AA.
 CC XX


```

146 9 MNSVSGESTSNLDECHVBLSTSNVNIWPSCHSLAHPYSSLLAHPGSSAWKAVAS 434
147
148 61 EHTLPVNETLAKKASNGASVTVTGCKRAHPAVSYVASGYHGWMSCECKKAF 120
149
150 69 EHTLPVNETLAKKASNGASVTVTGCKRAHPAVSYVASGYHGWMSCECKKAF 128
151
152 121 KESTLGHNYTGATWCTTDKRRKSCVACPLKCYVYVWVGYSSEPEYVETLVEK 169
153
154 129 KESTLGHNYTGATWCTTDKRRKSCVACPLKCYVYVWVGYSSEPEYVETLVEK 188
155
156 181 SADBGLHCACKARSGSHAVRELLLDALSTFQVLTLLAEPPHVLSPESAFETAS 240
157
158 189 SSBGLVHGTSKARKNCHAHKAVELLSTLSPGVLTLLEAPGVNLVSPSMPTFAS 248
159
160 241 MMSSTLRADKRELVIHMSAKKIDCVELSTFQVPLPSYGMVYMMIMPSSTDRCK 400
161
162 249 MMSSTLRADKRELVIHMSAKKIDCVELSTFQVPLPSYGMVYMMIMPSSTDRCK 408
163
164 401 LTFAPDLVDROBKVEGILLLEPDLALISERKELKDHREYLVAMILLNSMYPL 460
165
166 409 LTFAPDLVDROBKVEGILLLEPDLALISERKELKDHREYLVAMILLNSMYPL 468
167
168 461 VTATGADSSRKLAHLINAVTALVWVIAKSGTSSGQSGMRANLMLLSHVRH 414
169
170 469 ASADGADSSRKTHLILNAVTVAVVIAKSGTSSGQSGMRANLMLLSHVRH 422
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500

```


PE Tumours
 XX
 XX Example 2: Fig 2c: 89pp: English.
 XX
 XX The present sequence is alternatively spliced isoform ER beta-4
 of the complete oestrogen receptor beta (ER beta c). Murine ER
 beta c clone mbr beta 4 (see AAV24945) has been identified, which
 includes newly described exon 5b, which is located in the third
 binding domain of the receptor and is probably involved in receptor
 function. ER beta-4 nucleic acid was obtained from rat ovaries.
 It includes exon 5b but lacks exon 6, containing a frameshift and
 truncation of the protein 13 amino acids beyond the translated
 exon 5b. Also disclosed are mbr beta-1 (AAV98125) and mbr beta 2
 (AAV98126) isoforms, and human exon 5b (AAV24944). Based on these
 sequences, the invention provides methods for identifying agents
 that block or augment ER beta c mediated transcriptional regulation,
 methods to determine whether ER beta c or its isoforms is being
 expressed in tissues or cells, and methods of identifying and using
 agents that block the transcriptional regulation of genes by ER
 beta c or its isoforms, which in turn modulates other biological
 and pathological processes. Gene therapy expression of ER beta c
 is envisaged.

XX Sequence 414 AA:
 Query Match 68.0% Score 14997 DB 202 Length 414:
 best local similarity 87.0% Prod. No. 1 to 134:
 Matches 274 conservative 16; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNSVSTSNVNLKGGGPGQITSNVWLTGSHSLVVRGSHSLVYAEVQKSPWTEAST 60
 DB 73 MNSVSTSNVNLKGGVGLSTSNVWLTGSHSLVYAEVQKSPWTEAST 132
 QY 61 EHTLPVNRKELTKKRSQNPASVPTGDSKRIAHFVAVCSIVASGYHWGWSFGKAF 120
 DB 133 EHTLPVNRKELTKKRLSSSSASVPTSPNKRRIAHFVAVCSIVASGYHWGWSFGKAF 192
 QY 121 KRSLGGHNDYTGATNCTTDKRRKRSQACRLKRCYEWGWSGSRPRGCRYRVRQR 180
 DB 193 KRSLGGHNDYTGATNCTTDKRRKRSQACRLKRCYEWGWSGSRPRGCRYRVRQR 252
 QY 181 SADPQHCAGKARSGHAPRVRELLDMSPPQVLTLLFAEPHYVLLSPSPAPTEAS 240
 DB 253 SSSGVCHGCKAKRNGHAPRVRELLDMSPPQVLTLLFAEPHYVLLSPSPAPTEAS 312
 QY 241 MMSSTKRLAKRELVHMLSMARKTPGFVELSLPGVRLLESWMVEVLMGTLMRSIDHCK 300
 DB 313 MMSSTKRLAKRELVHMLSMARKTPGFVELSLPGVRLLESWMVEVLMGTLMRSIDHCK 372
 QY 401 LTFAPDVLDR 411
 DB 473 LTFAPDVLDR 483

RESULT 28
 AAY04436
 ID AAY04436 standard: Protein: 730 AA.
 AC AAY04436:
 XX
 XX 25 JUN 1999 (first entry)
 XX
 XX Rat ER beta 4 clone protein sequence.
 XX
 XX Estrogen receptor beta: oestrogen receptor beta: ER beta.
 XX
 XX Ratfus sp.
 XX
 XX W09905171 AL
 XX
 XX 04 FEB 1999.
 XX
 XX 28 JUL 1998; 98WO 0815540.

XX
 XX 40 JUL 1997; 9708 0054210.
 XX
 XX 28 JUL 1997; 9708 0054669.
 XX
 XX (OXYA) UNIV YALE.
 XX
 XX Barton R, Chabon B, Denton R, Dupont S, Gartner J:
 DB W11: 1999 142854/12.

PE Novel murine oestrogen receptor-beta genes and splice variants
 useful for treating oestrogen receptor-beta mediated disorders
 claim 9: Fig 2c: 89pp: English.

The present invention describes the murine complete oestrogen receptor
 beta (mbr beta c) also designated mbr beta-4). The present sequence
 represents the rat isoform ER beta 4 of mbr beta c. The protein can
 be used to screen for agents that modulate or block the interaction
 between the ER beta c protein and physiologically, stilbene or
 triphenylethylene (ant) oestrogens. The antibody which binds the protein
 can be used in assays to detect activation of the ER beta protein and
 can be used for subcellular localisation of the protein. The
 complementary nucleic acid can be used to inhibit expression of the
 ER beta c nucleic acid.

XX Sequence 740 AA:
 Query Match 68.0% Score 14997 DB 202 Length 740:
 best local similarity 87.0% Prod. No. 46-144:
 Matches 274 conservative 16; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNSVSTSNVNLKGGGPGQITSNVWLTGSHSLVVRGSHSLVYAEVQKSPWTEAST 60
 DB 73 MNSVSTSNVNLKGGVGLSTSNVWLTGSHSLVYAEVQKSPWTEAST 132
 QY 61 EHTLPVNRKELTKKRSQNPASVPTGDSKRIAHFVAVCSIVASGYHWGWSFGKAF 120
 DB 133 EHTLPVNRKELTKKRLSSSSASVPTSPNKRRIAHFVAVCSIVASGYHWGWSFGKAF 192
 QY 121 KRSLGGHNDYTGATNCTTDKRRKRSQACRLKRCYEWGWSGSRPRGCRYRVRQR 180
 DB 193 KRSLGGHNDYTGATNCTTDKRRKRSQACRLKRCYEWGWSGSRPRGCRYRVRQR 252
 QY 181 SADPQHCAGKARSGHAPRVRELLDMSPPQVLTLLFAEPHYVLLSPSPAPTEAS 240
 DB 253 SSSGVCHGCKAKRNGHAPRVRELLDMSPPQVLTLLFAEPHYVLLSPSPAPTEAS 312
 QY 241 MMSSTKRLAKRELVHMLSMARKTPGFVELSLPGVRLLESWMVEVLMGTLMRSIDHCK 300
 DB 313 MMSSTKRLAKRELVHMLSMARKTPGFVELSLPGVRLLESWMVEVLMGTLMRSIDHCK 372
 QY 401 LTFAPDVLDR 411
 DB 473 LTFAPDVLDR 483

RESULT 29
 AAG78227
 ID AAG78227 standard: Protein: 582 AA.
 AC AAG78227:
 XX
 XX 14 DEC 2001 (first entry)
 XX
 XX Leptomis centarchidae oestrogen receptor S60 (p No 4).
 XX
 XX Leptomis centarchidae oestrogen receptor.
 XX
 XX Leptomis centarchidae.
 XX
 XX 102001197890-A.

PD 24-JUL-2001.
 XX
 PF 08-NOV-2000: 2000JP-0340097.
 XX
 PK 09-NOV-1999: 99JP-0316113.
 XX
 PA (SDMO) SUMITOMO CHEM CO LTD.
 XX
 DR WPI: 2001-609402/70.
 XX
 DR N-PSDB: AA168126.
 XX
 PI New polypeptide for controlling oestrogen receptor activity comprises
 PT the oestrogen receptor gene -
 XX
 PS claim 1: Page 23-24; 34pp; Japanese.
 XX
 CC The invention relates to oestrogen receptors and the encoding genes
 CC of the oestrogen receptor genes can be used for evaluating the ability of
 CC controlling oestrogen receptor activity of a chemical substance.
 XX
 SO Sequence 582 AA:
 Query Match 48.8%; Score 1075.5; DB 22; Length 592.
 Best Local Similarity 50.7%; Pred. No. 6,9e-94;
 Matches 211; Conservative 65; Mismatches 95; Indels 65; Gaps 12;
 YY 4 YSLP-----SNVTNAGGPRGRTSPNVMWTPGHSPLVYHRLSHLYAEVOKSP 54
 DB 41 YSNVPLDAGPSPDNGSLGSGP---TSPLVYVSSRLSPPM-IPPSHHYLETSTP 94
 YY 54 WCFANSLHTLELVNREFTLKRKVSCKKFCASP-----VTGPGSK-----RQAHFVAVSD 101
 DB 95 VVRSVSPSSQGVPPH-----TGCATSESYSVGSGAGACGCEPAKEMRGTAVGSD 146
 YY 102 YASGTHYGVWCEGCKAFKFKRSIQGHNDYITPAINOCTIDNRKSCQAFKRCYVGM 161
 DB 147 YASGTHYGVWCEGCKAFKFKRSIQGHNDYITPAINOCTIDNRKSCQAFKRCYVGM 206
 YY 162 VRCGSPRGPGYRIYVWPSA-----DEQHGCKAKKRGSGAPR 201
 DB 207 MKGIVKRR-RVYIFETKRPATNTNIEFKASKI ERYIVPGGEPHSSSSACGCGSPS 265
 YY 202 VRELLDALSPDQIVLTLEAEPPVYLISRP--SAPFTASMMSTIKLAKELVIMISW 259
 DB 266 V-----TMSPPQVILLGGAPP-MTCSPTKISPTFTVITMILTSMAKELVIMISW 319
 YY 260 AKKIPGFVSLFDVRIIESGVVIMMGLMBSIDHPKILFAPDLVIDEKKVEG 319
 DB 320 AKKIPGFVSLFDVRIIESGVVIMMGLMBSIDHPKILFAPDLVIDEKKVEG 379
 YY 320 IIFETFMILATSPFEFKIQHFEYCVKAMILLNSSMYPLVATGQ-ALSSPELAILN 378
 DB 380 FVIEITMLLATASRRHMLKIKPEFVCIKAILLINSAPSTGTGIMPLINSMAVQNMID 439
 YY 379 AVTIDALVWIAKSGISSQOOSRLANLLMLLSVHR 414
 DB 440 TITDALIHHSQSGSSAQOOSRRQADILLISHIRH 475
 RESULT 30
 ABB81600
 ID ABB81600 standard: Protein: 587 AA.
 XX
 AC ABB81600:
 XX
 DT 19-SEP-2002 (first entry)
 XX
 DE Human crocodilius oestrogen receptor protein SEQ ID NO:1.
 XX
 KW Human crocodilius: oestrogen receptor; endocrine disrupter, receptor.
 XX
 OS Human crocodilius.
 XX

PN W0200248360-A1.
 XX
 PD 29 JUN 2002.
 XX
 PK 15-NOV-2001: 2001WO-JP09996.
 XX
 PA 14-06-2000: 2000JP-0379993.
 XX
 DR WPI: 2002-557614/59.
 XX
 DR N-PSDB: AB072931.
 XX
 PI Sumida K:
 XX
 PI Crocodile-orientated estrogen receptor genes, applicable in test
 PT systems for evaluating the ability of chemicals to control estrogen
 PT receptor activity e.g. to assay the safety of chemicals and detecting
 PT endocrine disrupters in environment
 XX
 PS claim 1: Page 60-63; 81pp; Japanese.
 XX
 CC The present sequence represents an oestrogen receptor isolated from
 CC human crocodilius. The oestrogen receptor gene can be used in testing
 CC systems for evaluating the ability of chemicals to control the
 CC oestrogen receptor activity, e.g. assay safety of chemicals and
 CC detecting endocrine disrupters in environment and the genes and the
 CC DNA are useful in two-hybrid assays.
 XX
 SO Sequence 587 AA:
 Query Match 48.8%; Score 1065; DB 23; Length 587;
 Best Local Similarity 52.7%; Pred. No. 7,1e-93;
 Matches 218; Conservative 69; Mismatches 105; Indels 24; Gaps 8;
 YY 22 SPNVLPPTGHSPLVYHRLSHLYAEVOKSPW--CFARSLHTLIPVNEFTLKRKYSN 78
 DB 100 SPVPLDAGPSPDNGSLGSGP---TSPLVYVSSRLSPPM-IPPSHHYLETSTP 156
 YY 79 WCFANSLHTLELVNREFTLKRKVSCKKFCASP-----VTGPGSK-----RQAHFVAVSD 133
 DB 157 ERMSSTSEKASLSESTKEPVCACVNDYASGVWCEGCKAFKFKRSIQGHNDYICP 216
 YY 134 ATNOCTIDNRKSCQAFKFKRSIQGHNDYITPAINOCTIDNRKSCQAFKFKRSIQGHNDYICP 193
 DB 217 ATNOCTIDNRKSCQAFKFKRSIQGHNDYITPAINOCTIDNRKSCQAFKFKRSIQGHNDYICP 276
 YY 194 RSGGAPPV-RELLD-----ALSPQIVLTLEAEPPVYLISRP-SAPFTASMM 241
 DB 277 TAEIRTPILTMSIVYIKHFKKNSPALSTAEQMSALILEAPPVYVSEYDNRPFNEASM 336
 YY 242 MKSITLAKRFLVHMSAKKIPGFVSLFDVRIIESGVVIMMGLMBSIDHPKILFAPDLVIDEKKVEG 301
 DB 437 MTITNLAPELVHIMNAKRVGPDVITINQVHIFCAVLPITMIGVPSMFWHDKL 496
 YY 402 IFAPDLVIDEKKVEGILFEMILATSPFEFKIQHFEYCVKAMILLNSSMYPLV 461
 DB 497 IFAPDLVIDEKKVEGILFEMILATSPFEFKIQHFEYCVKAMILLNSSMYPLV 456
 YY 462 TATUDANSSKLAH-ILNAVTDALVWIAKSGISSQOOSRLANLLMLLSVHR 414
 DB 457 SSTLKSSPEKIVAVVIAKTIITITIMAKSSSTSGGSHRPAADITISHIRH 510
 RESULT 31
 AAG84509
 ID AAG84509 standard: Protein: 595 AA.
 XX
 AC AAG84509:
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human oestrogen receptor alpha protein mutant M99V.
 XX

XX 03-JUN-1999.
 XX
 XX 25-NOV-1998; 98W0 0525296.
 XX
 XX 26-NOV-1997; 97W0 0980115.
 XX
 XX (RESC) UNIV CALIFORNIA.
 XX
 XX Appt 1011 JW, Baxter JB, Fletcher KJ, Kushner EJ,
 XI Scandan TS, Shiao AK, Warner RL, West BL;
 XX
 XX Wnt: 1999 457410/70.
 XX
 XX Modulating activity of a thyroid hormone receptor
 XI
 XX
 XX Disclosure: Fig 4G K; 44pp; English.
 XX
 XX The invention relates to a method for modulation activity of a thyroid
 XX hormone receptor that comprises administration of an aromatic compound
 XX which is spatially and preferentially into a thyroid hormone ligand
 XX binding domain. The aromatic compound (of a specified formula) can be
 XX used to increase alpha triphosphate dehydrogenase (GPDH) levels, at
 XX levels which do not significantly modify cardiac GPDH levels and are
 XX indicated in the treatment of obesity. The compound also lower total
 XX plasma cholesterol and triacylglyceride levels and can be used as anti
 XX hypertriglyceridemic agents. The compound may also be used for treating
 XX atherosclerosis and may be indicated in thyroid hormone replacement
 XX therapy in patients with compromised cardiac function. Sequences
 XX AAY21621 646 amino acid sequences of ligand binding domains of several
 XX members of the nuclear receptor superfamily.
 XX
 XX Sequence: 595 AA;
 XX
 XX Query Match: 48-48; Score 1063; 108 22; Length 595;
 XI Host Local Similarity: 48-78; Prod. No. 1, Loc 92;
 XI Matches: 273; Conservative: 78; Mismatches: 99; Indels: 54; Gaps: 10;
 XX
 XX 10 TNEKRGKQ-----TTSNVMPTPGHSP LVVHQSLSHYAVQ 50
 XI 1
 XI 76 TGLPYGNSGAAPGNSNGLGPPPLNSVSPSLMLHPPLQSLPGHGGVYLYENE 145
 XX
 XX 51 KSIWQ- EAKSLHETLPVNRKELTKKVSINRQASVYDQ-----SKDAHPVAVSIVA 104
 XI 1
 XI 146 PSQVYREKNGPPAVPR- NSDNKQGGREKLASTNKGSAWMSAKETPVAVCNVA 192
 XX
 XX 104 STHYGVWSTEG- KAFKESIGSHNYICPATINQTIORPKPSQVGLKRYEVSQK 163
 XI 1
 XI 194 STHYGVWSTEG- KAFKESIGSHNYICPATINQTIORPKPSQVGLKRYEVSQK 252
 XX
 XX 164 GNSRRERGVRLVRRQSSMDQDLRACAKKRSQADPRV-----ELL 207
 XI 1
 XI 254 GRIKRDGKGMKIKRQKRD-----GGRGRGVSSAMRANLMSPLMKRSKNSIA 307
 XX
 XX 208 DALSCELVLTLEAPPHVLISR PSAPTEANMMMSIKRARELVIHNSMAKELPV 255
 XI 1
 XI 408 LSTLALQVNSALDAVPR- ILYSPQPIPPSPSPAMGRIINLADELVHNMNAKRVPS 466
 XX
 XX 266 FVSLSTPVQVRLLESTQMEVLMQLMWSTIDHPEKELFAPDVLTRDQEK-VEGLLEFD 325
 XI 1
 XI 467 FVSLSTPVQVRLLESTQMEVLMQLMWSTIDHPEKELFAPDVLTRDQEK-VEGLLEFD 426
 XX
 XX 426 MLATTSREPEKELQHEVLEVRADLINSMTYLVAT QVANSKRLALMLAAVTHAL 484
 XI 1
 XI 427 MLATTSREPEKELQHEVLEVRADLINSMTYLVAT QVANSKRLALMLAAVTHAL 490
 XX
 XX 485 VWVIAKSGTSSQGSQMLANLMLSHYH 414
 XI 1
 XI 487 TTMMAKAGLTLDQDQHQDGLADLILSHIRH 516
 XX

RESULT 35
 AAB84505

10 AAB84505, standard; Protein; 595 AA.
 XX
 XX AC AAB84505;
 XX
 XX 10 SEP-2001 (first entry)
 XX
 XX Human oestrogen receptor alpha protein.
 XX
 XX
 XX Liquid dependent transcriptional factor; oestrogen receptor; ER;
 XX glucocorticoid receptor protein; GR; mineralocorticoid receptor protein;
 XX MR; peroxisome proliferator activated receptor protein; PPAR;
 XX progesterone receptor protein; PR; progamine X receptor protein; PXR;
 XX thyroid hormone receptor protein; TR; vitamin D receptor protein; VDR;
 XX transactivation; Eralpha; breast cancer.
 XX
 XX Homo sapiens.
 XX
 XX W0200142407 A1.
 XX
 XX 14 JUN-2001.
 XX
 XX 01 DEB-2000; 2000W0-JP08553.
 XX
 XX 07 DEB-1999; 99JP 0348922.
 XX
 XX 27 DEB-1999; 99JP 0370667.
 XX
 XX 07 JUN-2000; 2000JP 0207011.
 XX
 XX 21 JUN-2000; 2000JP 0220508.
 XX
 XX 02 AUG-2000; 2000JP 0244053.
 XX
 XX 03 AUG-2000; 2000JP 0235460.
 XX
 XX 03 AUG-2000; 2000JP 0235461.
 XX
 XX 03 AUG-2000; 2000JP 0235463.
 XX
 XX (SOMO) SUMITOMO CHEM CO LTD.
 XX
 XX Saito K, Ohe N, Saito H;
 XI
 XX W01: 2001-067866/18.
 XX
 XX Liquid dependent transcriptional factors, nucleic acids encoding them
 XI and cells comprising them and a specified reporter gene, useful for
 XI screening agents for the treatment of breast cancer.
 XI
 XI Claim 7: Page 167-170; 276pp; English.
 XI
 XI The present invention relates to ligand dependent transcriptional factors
 XI including oestrogen receptor (ER) alpha and beta protein, glucocorticoid
 XI receptor protein (GR), mineralocorticoid receptor protein (MR),
 XI peroxisome proliferator activated receptor protein (PPAR), progesterone
 XI receptor protein (PR), progamine X receptor protein (PXR), thyroid hormone
 XI receptor protein (TR) and vitamin D receptor protein (VDR). The nucleic
 XI acids encoding them and cells comprising them and a specified reporter
 XI gene for the ligand dependent transcriptional factor. These proteins are
 XI useful in the modulation of ligand dependent transcriptional factor
 XI activity. The cells, mutant Eralpha and the polynucleotide encoding it
 XI may be used in assays for qualitatively analyzing an activity for
 XI transactivation of a reporter gene by a test Eralpha, for screening
 XI mutant ligand dependent transcriptional factors, for evaluating an
 XI activity for transactivation of a reporter gene by a test Eralpha and/or
 XI for screening a compound useful for treating a disorder of a mutant
 XI Eralpha, especially breast cancer.
 XX
 XX Sequence: 595 AA;
 XX
 XX Query Match: 48-48; Score 1063; 108 22; Length 595;
 XI Host Local Similarity: 48-78; Prod. No. 1, Loc 92;
 XI Matches: 273; Conservative: 78; Mismatches: 99; Indels: 54; Gaps: 10;
 XX
 XX 10 TNEKRGKQ-----TTSNVMPTPGHSP LVVHQSLSHYAVQ 50
 XI 1
 XI 76 TGLPYGNSGAAPGNSNGLGPPPLNSVSPSLMLHPPLQSLPGHGGVYLYENE 145
 XX
 XX 51 KSIWQ- EAKSLHETLPVNRKELTKKVSINRQASVYDQ-----SKDAHPVAVSIVA 104
 XI 1

```

Db 136 PGGYIVREAGPPAFYRP---NSDNPRGGGRRRLASTNRKQSMAMSAKETPVAVNINVA 192
QY 104 SGRHYGVMS-EG-KAFKPSLQ-SHNYVLPATNR-ITTFNFKS-DAV-RELPV-YEVNVA 163
Db 193 SGRHYGVMS-EG-KAFKPSLQ-SHNYVLPATNR-ITTFNFKS-DAV-RELPV-YEVNVA 252
QY 164 GTSPPRPGVPIVHPGNSADPQHPACAKAPSGGAPVPV-----GILL 207
Db 253 GTSPPRPGVPIVHPGNSADPQHPACAKAPSGGAPVPV-----GILL 297
QY 208 DMLSPQVITLIFAPRHVILSP--PSAFETFAAMMSITKLADKELVHMSAKETFG 265
Db 308 LSLTAQVMSALLDAAPP-ILYSEYDPIRPSFAAMGILLNLADRELVHMINNAKVPFG 366
QY 266 EVELSTFQVPLIFSRVGVLMWGIWMSITRPECLITADQVIDELFKVVEGILLTFG 325
Db 467 FVDLTHDQVHLIFAMLEIMIGLWRSMEHPKELFAPNILLDRNGKVEGIVLEFD 426
QY 426 MLATTSRRELKLOHKEVLCVAKMILLNSMYPLVTAI-QDAUSSRKLAMILNAVTAL 484
Db 427 MLATTSRPRMNIQGEFFVCEKSTILLNSVYTLSTLSKEERKIHIVLKKITDTL 494
QY 485 VWYIAKSGISSQGSMLANLMLLSHVKH 414
Db 487 IILMAKAGLTLQDQHQRALQILLLSHRH 516

```

RESULT 36

AA084506 standard; Protein: 595 AA.

AA084506:

10-SEP-2001 (first entry)

Human oestrogen receptor alpha protein mutant K303P.

Ligand dependent transcriptional factor: oestrogen receptor; ER.

glucocorticoid receptor protein; GR; mineralocorticoid receptor protein;

MR; peroxisome proliferator-activated receptor protein; PPAR;

progestrone receptor protein; PR; progesterone X receptor protein; PRX;

thyroid hormone receptor protein; TR; vitamin D receptor protein; VDR;

transactivation; ERalpha; breast cancer; mutant; mutlen.

Homo sapiens.

W0200142107-A1.

07-DEC-1999: 99JP-0348022.

07-DEC-1999: 99JP-0370667.

07-JUN-2000: 2000JP-0207011.

02-AUG-2000: 2000JP-0234493.

03-AUG-2000: 2000JP-0235460.

03-AUG-2000: 2000JP-0235461.

03-AUG-2000: 2000JP-0235463.

(SOMO) SOMITOMO CHEM CO LTD.

Saito K, Oho N, Satoh H.

WPI: 2001-367866/38.

the present invention relates to ligand dependent transcriptional factors including oestrogen receptor (ER) alpha and beta protein, glucocorticoid receptor protein (GR), mineralocorticoid receptor protein (MR), peroxisome proliferator-activated receptor protein (PPAR), progesterone receptor protein (PR), progesterone X receptor protein (PRX), thyroid hormone receptor protein (TR) and vitamin D receptor protein (VDR). The nucleic acids encoding them and cells comprising them and a specified reporter gene for the ligand dependent transcriptional factor. These proteins are useful in the modulation of ligand dependent transcriptional factor activity. The cells, mutant ERalpha and the polynucleotide encoding it may be used in assays for qualitatively analysing an activity for CC transactivation of a reporter gene by a test ERalpha, for screening CC mutant ligand dependent transcriptional factors, for evaluating an activity for transactivation of a reporter gene by a test ERalpha and/or for screening a compound useful for treating a disorder of a mutant ERalpha, especially breast cancer.

Sequence 595 AA:

Query Match 48.3%; Score 1063; DR 22; Length 595;

Best Local Similarity 48.7%; Pred. No. 1.3e-927

Matches 219; Conservative 78; Mismatches 99; Indels 54; Gaps 10;

```

QY 10 INIPGPGG-----TTSNVIWPPRHISPIVVPQISPIYAPPG 50
Db 76 TGIPGPGGSAAPGNSNIGGPPILNSVSPSILMLHPPIQISPIYGVYIENE 145
QY 51 KSPW--EASLHLLVNBETLKKVSNVCASTVTCG SQAHHVAVNSIVA 103
Db 136 PGGYIVREAGPPAFYRP---NSDNPRGGGRRRLASTNRKQSMAMSAKETPVAVNINVA 192
QY 104 SGRHYGVMS-EG-KAFKPSLQ-SHNYVLPATNR-ITTFNFKS-DAV-RELPV-YEVNVA 163
Db 193 SGRHYGVMS-EG-KAFKPSLQ-SHNYVLPATNR-ITTFNFKS-DAV-RELPV-YEVNVA 252
QY 164 GTSPPRPGVPIVHPGNSADPQHPACAKAPSGGAPVPV-----GILL 207
Db 253 GTSPPRPGVPIVHPGNSADPQHPACAKAPSGGAPVPV-----GILL 297
QY 208 DMLSPQVITLIFAPRHVILSP--PSAFETFAAMMSITKLADKELVHMSAKETFG 265
Db 308 LSLTAQVMSALLDAAPP-ILYSEYDPIRPSFAAMGILLNLADRELVHMINNAKVPFG 366
QY 266 EVELSTFQVPLIFSRVGVLMWGIWMSITRPECLITADQVIDELFKVVEGILLTFG 325
Db 467 FVDLTHDQVHLIFAMLEIMIGLWRSMEHPKELFAPNILLDRNGKVEGIVLEFD 426
QY 426 MLATTSRRELKLOHKEVLCVAKMILLNSMYPLVTAI-QDAUSSRKLAMILNAVTAL 484
Db 427 MLATTSRPRMNIQGEFFVCEKSTILLNSVYTLSTLSKEERKIHIVLKKITDTL 494
QY 485 VWYIAKSGISSQGSMLANLMLLSHVKH 414
Db 487 IILMAKAGLTLQDQHQRALQILLLSHRH 516

```

RESULT 37

AA084513 standard; Protein: 595 AA.

AA084513:

10-SEP-2001 (first entry)

Human oestrogen receptor alpha protein mutant S578P.

Ligand dependent transcriptional factor: oestrogen receptor; ER.

glucocorticoid receptor protein; GR; mineralocorticoid receptor protein;

MR; peroxisome proliferator-activated receptor protein; PPAR;

progestrone receptor protein; PR; progesterone X receptor protein; PRX;

thyroid hormone receptor protein; TR; vitamin D receptor protein; VDR;

transactivation; ERalpha; breast cancer; mutant; mutlen.

101 342 15.1 446 5 pct US95-16311 3
 102 341.5 15.0 443 4 ES-08-442 411A 4
 103 340 15.0 446 4 US-08-776 84A 2
 104 327.5 14.9 460 1 US-08-342 411A 2
 105 325.5 14.6 672 6 5171671-2
 106 320.5 14.5 674 4 US-08-653-64HA-14
 107 320.5 14.5 764 3 US-09-144-759 20
 108 317.5 14.4 448 6 US-08-653-64HA-15
 109 317.5 14.4 878 4 US-08-764 870A-15
 110 317 14.4 416 4 US-08-764 870A-15
 111 317 14.4 416 4 US-08-980-115 4
 112 317 14.4 797 2 US-08-995 728B 2
 113 317 14.4 797 5 pct US92 02320A-2
 114 316 14.3 368 6 5223606-3
 115 316 14.3 575 4 US-08-653-64HA-5
 116 312.5 14.2 746 3 US-09-144-759-18
 117 307.5 14.0 410 4 US-08-764-870-2
 118 307.5 14.0 410 4 US-08-980-115-2
 119 307.5 14.0 410 6 548126-2
 120 306 13.9 557 4 US-08-653-64HA-12

ALIGNMENTS

Sequence 3, Appl1
 Sequence 4, Appl1
 Sequence 2, Appl1
 Patient No. 5171671
 Sequence 14, Appl1
 Sequence 20, Appl1
 Patient No. 5223606
 Sequence 15, Appl1
 Sequence 4, Appl1
 Sequence 2, Appl1
 Patient No. 5223606
 Sequence 5, Appl1
 Sequence 18, Appl1
 Sequence 2, Appl1
 Patient No. 548126
 Sequence 12, Appl1

RESULT 1
 US-08-826-620A 4
 Sequence 3, Application US-08-826-620A
 Patient No. 5958710
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: orphan receptor
 NUMBER OF SEQUENCES: 19
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: Patient In keyboard #1.0, Version #1.25 (EPR)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US-08-826-620A
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP96/03933
 FILING DATE:
 APPLICATION NUMBER: GB 9518272.1
 FILING DATE: 08 SEP 1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9605550.4
 FILING DATE: 15 MAR 1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9607542.0
 FILING DATE: 11 APR 1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9609576.5
 FILING DATE: 08 MAY 1996
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 485 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGIN: SOURCE:
 ORGANISM: Homo sapiens
 US-08-826-620A 4
 Query Match 99.8%, Score 2198, Dh 2, Length 485
 Post Local Similarity 100.0%, Pred. No. 1, No-244
 Matches 415, Conservative 0, Mismatches 0, Indels 0, Gaps 0
 1 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 60
 61 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 60
 9 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 60
 61 EHTLPVNRRLTKRKVSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 120

|||||
 69 EHTLPVNRRLTKRKVSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 120
 121 KSTIGRNINVTVAIWSCTTGNKPKSVAACILPKCYEVMSKSSPPECCYPIVPPG 180
 129 KSTIGRNINVTVAIWSCTTGNKPKSVAACILPKCYEVMSKSSPPECCYPIVPPG 180
 181 SAEQDGLAKAKRSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 240
 189 SAEQDGLAKAKRSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 248
 241 KSTIGRNINVTVAIWSCTTGNKPKSVAACILPKCYEVMSKSSPPECCYPIVPPG 300
 249 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 308
 301 LTPAPDLVLDROGKCVESGLETFDMLATTSRPFELKQKREYLVKAMILLNSMYPL 360
 309 LTPAPDLVLDROGKCVESGLETFDMLATTSRPFELKQKREYLVKAMILLNSMYPL 368
 361 VTATQADSSKRLAHLLNAVVTALVWVIAKSGISSQOOSMRILMLLSIVRIA 415
 469 VTATQADSSKRLAHLLNAVVTALVWVIAKSGISSQOOSMRILMLLSIVRIA 423

RESULT 2
 US-09-149-617-1
 Sequence 1, Application US-09-149-617
 Patient No. 622015
 GENERAL INFORMATION:
 APPLICANT: WILKINSON, HILARY
 TITLE OF INVENTION: ESTROGEN RECEPTOR
 FILE REFERENCE: 20047Y
 CURRENT APPLICATION NUMBER: US-09-149-617
 CURRENT FILING DATE: 1998-08-25
 EARLIER APPLICATION NUMBER: 60/058,271
 EARLIER FILING DATE: 1997-09-08
 EARLIER APPLICATION NUMBER: 60/060,520
 EARLIER FILING DATE: 1997-09-10
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: Fast Seq for Windows Version 4.0
 SEQ ID NO: 1
 LENGTH: 548
 TYPE: PRT
 ORGANISM: HUMAN
 US-09-149-617-1
 Query Match 99.8%, Score 2198, Dh 4, Length 548
 Post Local Similarity 100.0%, Pred. No. 1, No-244
 Matches 415, Conservative 0, Mismatches 0, Indels 0, Gaps 0
 1 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 60
 72 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 131
 61 EHTLPVNRRLTKRKVSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 120
 132 EHTLPVNRRLTKRKVSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 191
 121 KSTIGRNINVTVAIWSCTTGNKPKSVAACILPKCYEVMSKSSPPECCYPIVPPG 180
 192 KSTIGRNINVTVAIWSCTTGNKPKSVAACILPKCYEVMSKSSPPECCYPIVPPG 251
 181 SAEQDGLAKAKRSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 240
 252 SAEQDGLAKAKRSNGASVPTGSGKRDAHRCVAVSYGCVMSCTGCKKAF 311
 241 KSTIGRNINVTVAIWSCTTGNKPKSVAACILPKCYEVMSKSSPPECCYPIVPPG 300
 412 MNSTISNVNLEKGGKGTTSNVLPKISPLVVRGQSHYATFGKSPWFAASL 371
 401 LTPAPDLVLDROGKCVESGLETFDMLATTSRPFELKQKREYLVKAMILLNSMYPL 360
 472 LTPAPDLVLDROGKCVESGLETFDMLATTSRPFELKQKREYLVKAMILLNSMYPL 431

UY 461 VTATGATGSSKRLAHILNAVITALLVWVIAKSGTSSQSSMPANILMLTSHVRA 415
 432 VTATGATGSSKRLAHILNAVITALLVWVIAKSGTSSQSSMPANILMLTSHVRA 486

RESULT 4 US-09-561-741A-1

Sequence 1, Application US/09561741A
 Patent No. 6458551
 GENERAL INFORMATION:
 APPLICANT: WILKINSON, HILARY
 TITLE OF INVENTION: ESTROGEN RECEPTOR
 FILING DATE: 20047Y
 CURRENT APPLICATION NUMBER: US/09/561,741A
 PRIOR APPLICATION NUMBER: 2000-04-26
 PRIOR FILING DATE: 09/13/99, 617
 PRIOR FILING DATE: 1998-08-25
 PRIOR APPLICATION NUMBER: 60/058,271
 PRIOR FILING DATE: 1997-09-08
 PRIOR APPLICATION NUMBER: 60/060,520
 PRIOR FILING DATE: 1997-09-30
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO: 1
 LENGTH: 548
 TYPE: PRT
 ORGANISM: HUMAN
 US-09-561-741A-1

Query Match 99.84; Score 2198; PR 4; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1,86,234;

Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 MNTSIPSNVNLDSGSGHUTTSFNVLPPTGRI SP1VHRQLSHLYAEQKSPWCEAFASL 60
 DB 72 MNTSIPSNVNLDSGSGHUTTSFNVLPPTGRI SHLYAEQKSPWCEAFASL 131
 UY 61 EHTLPVNRKTLKRVKSNRCPASVPTGSGKRDHAFVAGSDVASYHNYVMSCEGKAF 120
 DB 132 EHTLPVNRKTLKRVKSNRCPASVPTGSGKRDHAFVAGSDVASYHNYVMSCEGKAF 191
 UY 121 KRSTIGHNDYICPATINOTIDKNNRKSQACRLKCYEVMWVQDSREPCGYPLVRROR 180
 DB 192 KRSTIGHNDYICPATINOTIDKNNRKSQACRLKCYEVMWVQDSREPCGYPLVRROR 251
 UY 181 SADBQLHCACAKRSGSHAPRVELLDALSPQLVTLLEAPRVLISRTSAPRTAS 240
 DB 252 SADBQLHCACAKRSGSHAPRVELLDALSPQLVTLLEAPRVLISRTSAPRTAS 311
 UY 241 MMTSITKLADKELVHMTSMARKTPGVEVLSLPGVRLLESQWMEVLMGLMPSITRQK 300
 DB 412 MMTSITKLADKELVHMTSMARKTPGVEVLSLPGVRLLESQWMEVLMGLMPSITRQK 371
 UY 301 LIFAPDLVLRDEQKCEVGLLETFDMLATTSFRELKQHKRYLCVKAMILLNSMYP 360
 DB 472 LIFAPDLVLRDEQKCEVGLLETFDMLATTSFRELKQHKRYLCVKAMILLNSMYP 431
 UY 461 VTATGATGSSKRLAHILNAVITALLVWVIAKSGTSSQSSMPANILMLTSHVRA 415
 DB 432 VTATGATGSSKRLAHILNAVITALLVWVIAKSGTSSQSSMPANILMLTSHVRA 486

RESULT 4
 US-08-836-620A-15
 Sequence 15, Application US/08846620A
 Patent No. 5958710
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Orphan receptor
 NUMBER OF SEQUENCES: 19
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-WINDOWS-95
 SOFTWARE: Patent In Release #1.0, Version #1.25 (EPI)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/836,620A
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP96/03933
 FILING DATE:
 APPLICATION NUMBER: GR 9518272.1
 FILING DATE: 08-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GR 9605550.4
 FILING DATE: 15-MAR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GR 9607532.0
 FILING DATE: 11-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GR 9609576.5
 FILING DATE: 08-MAY-1996
 INFORMATION FOR SEQ ID NOS: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 384 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 US-08-836-620A-15

Query Match 91.64; Score 2017; PR 2; Length 384;
 Best Local Similarity 100.0%; Pred. No. 1,16,214;
 Matches 383; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

UY 33 LSP1VHRQI SHLYAEQKSPWCEAFASL EHTLPVNRKTLKRVKSNRCPASVPTGSGKRD 92
 DB 2 LSP1VHRQI SHLYAEQKSPWCEAFASL EHTLPVNRKTLKRVKSNRCPASVPTGSGKRD 61
 UY 93 AHTFVAGSDVASYHNYVMSCEGKAFKRSIGDHNYICPATINOTIDKNNRKSQACRL 152
 DB 62 AHTFVAGSDVASYHNYVMSCEGKAFKRSIGDHNYICPATINOTIDKNNRKSQACRL 121
 UY 153 LKCYEVMWVQDSREPCGYPLVRROR SADBQLHCACAKRSGSHAPRVELLDALSP 212
 DB 122 LKCYEVMWVQDSREPCGYPLVRROR SADBQLHCACAKRSGSHAPRVELLDALSP 181
 UY 213 KQVLTLEAPRVLISRTSAPRTASMMMSITKLADKELVHMTSMARKTPGVEVLSL 272
 DB 182 KQVLTLEAPRVLISRTSAPRTASMMMSITKLADKELVHMTSMARKTPGVEVLSL 241
 UY 273 DQVRLLESQWMEVLMGLMPSITRQK LIFAPDLVLRDEQKCEVGLLETFDMLATTS 332
 DB 242 DQVRLLESQWMEVLMGLMPSITRQK LIFAPDLVLRDEQKCEVGLLETFDMLATTS 301
 UY 333 RFRRLKQHKRYLCVKAMILLNSMYPVTATGATGSSKRLAHILNAVITALLVWVIAKSG 392
 DB 302 RFRRLKQHKRYLCVKAMILLNSMYPVTATGATGSSKRLAHILNAVITALLVWVIAKSG 361
 UY 493 ISSQSSMPANILMLTSHVRA 415
 DB 462 ISSQSSMPANILMLTSHVRA 486

RESULT 5
 US-08-836-620A-13
 Sequence 13, Application US/08846620A
 Patent No. 5958710
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: Orphan receptor
 NUMBER OF SEQUENCES: 19
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible


```

CURRENT APPLICATION DATA:
1  APPLICATION NUMBER: US/08/836,620A
2  FILING DATE:
3  PRIOR APPLICATION DATA:
4  APPLICATION NUMBER: PCT/EP96/03933
5  FILING DATE:
6  APPLICATION NUMBER: GB 9518272.1
7  FILING DATE: 08-SEP-1995
8  PRIOR APPLICATION DATA:
9  APPLICATION NUMBER: GB 9605550.4
10 FILING DATE: 15-MAR-1996
11 PRIOR APPLICATION DATA:
12 APPLICATION NUMBER: GB 9607532.0
13 FILING DATE: 11-APR-1996
14 PRIOR APPLICATION DATA:
15 APPLICATION NUMBER: GB 9609576.5
16 FILING DATE: 08-MAY-1996
17 INFORMATION FOR SEQ ID NO: 14:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 484 amino acids
20 TYPE: amino acid
21 TOPOLOGY: linear
22 ORIGINAL SOURCE:
23 ORGANISM: Mus musculus
24 US-08-836-620A-14

```

```

Query Match: 89.28; Score 1965; DB 2; Length 484;
Best Local Similarity: 88.68; Prod. No. 9.4e-209;
Matches 367; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

```

```

1  MNSIRSNVNLGCGRGQRTSDNVIWPTPHLSPIVWHPQLSHYAFQKSPWCFAPSI 60
2  IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
3  9  MNSVSSSTGNLEGGVROTASPNVLMPTSGHLSPLATHCQSSLLTAEPQKSPWCFAPSI 68
4  DB
5  QY 61  EHTLPVNPETLKPKVSGNRCASPVTPGSKRAHFCVAVCSYASGYHGVWSCGCAAF 120
6  IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
7  DB 69  EHTLPVNPETLKPKVSGNRCASPVTPGSKRAHFCVAVCSYASGYHGVWSCGCAAF 128
8  QY 121 KRSTGCHNDYICPATNCTIDKNPRKQACPLFKCYGVGWKCGSRHFGYRIVVRQK 180
9  IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
10 DB 129 KRSTGCHNDYICPATNCTIDKNPRKQACPLFKCYGVGWKCGSRHFGYRIVVRQK 188
11 QY 181 SAEQQLHCGAKKRSQGHAPRVPELLDALSPEDVLTLLAEPPHVLISRPASPTFAS 240
12 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
13 DB 189 SAEQYHCLNKAKRTSGHTPRVKELLNLSPEQVLTLLAEPPHVLISRPASPTFAS 248
14 QY 241 MMSSTLRKADKELVHMTSAKKIPGVVELSLFDQYRLLESCEMVEYLMKGLMKRSIDHFGK 300
15 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
16 DB 249 MMSSTLRKADKELVHMTSAKKIPGVVELSLFDQYRLLESCEMVEYLMKGLMKRSIDHFGK 308
17 QY 301 LIFAPQVLDROEGKVEGILELFFMLLATTSRPELKIQHKRYLCVAMILLNSMPTL 360
18 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
19 DB 309 LIFAPQVLDROEGKVEGILELFFMLLATTSRPELKIQHKRYLCVAMILLNSMPTL 368
20 QY 361 VTATQDQSSPKLAHLINAVTDLVWYAKSGTSSQGSMPANILMLLSHVRH 414
21 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
22 DB 369 ATASQAEASSPKLTHIINAVTDLVWYAKSGTSSQGSMPANILMLLSHVRH 422

```

```

RESULT 8
US-08-836-620A-5
1 Sequence 5; Application US/08/836,620A
2 Patient No. 5958710
3 GENERAL INFORMATION:
4 APPLICANT:
5 TITLE OF INVENTION: Orphan receptor
6 NUMBER OF SEQUENCES: 19
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
12 CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/836,620A
1  FILING DATE:
2  PRIOR APPLICATION DATA:
3  APPLICATION NUMBER: PCT/EP96/03933
4  FILING DATE:
5  APPLICATION NUMBER: GB 9518272.1
6  FILING DATE: 08-SEP-1995
7  PRIOR APPLICATION DATA:
8  APPLICATION NUMBER: GB 9605550.4
9  FILING DATE: 15-MAR-1996
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: GB 9607532.0
12 FILING DATE: 11-APR-1996
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: GB 9609576.5
15 FILING DATE: 08-MAY-1996
16 INFORMATION FOR SEQ ID NO: 5:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 485 amino acids
19 TYPE: amino acid
20 TOPOLOGY: linear
21 ORIGINAL SOURCE:
22 ORGANISM: Mus musculus
23 US-08-836-620A-5

```

```

Query Match: 89.28; Score 1965; DB 2; Length 485;
Best Local Similarity: 88.68; Prod. No. 9.4e-209;
Matches 367; Conservative 21; Mismatches 26; Indels 0; Gaps 0;

```

```

1  MNSIRSNVNLGCGRGQRTSDNVIWPTPHLSPIVWHPQLSHYAFQKSPWCFAPSI 60
2  IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
3  9  MNSVSSSTGNLEGGVROTASPNVLMPTSGHLSPLATHCQSSLLTAEPQKSPWCFAPSI 68
4  DB
5  QY 61  EHTLPVNPETLKPKVSGNRCASPVTPGSKRAHFCVAVCSYASGYHGVWSCGCAAF 120
6  IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
7  DB 69  EHTLPVNPETLKPKVSGNRCASPVTPGSKRAHFCVAVCSYASGYHGVWSCGCAAF 128
8  QY 121 KRSTGCHNDYICPATNCTIDKNPRKQACPLFKCYGVGWKCGSRHFGYRIVVRQK 180
9  IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
10 DB 129 KRSTGCHNDYICPATNCTIDKNPRKQACPLFKCYGVGWKCGSRHFGYRIVVRQK 188
11 QY 181 SAEQQLHCGAKKRSQGHAPRVPELLDALSPEDVLTLLAEPPHVLISRPASPTFAS 240
12 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
13 DB 189 SAEQYHCLNKAKRTSGHTPRVKELLNLSPEQVLTLLAEPPHVLISRPASPTFAS 248
14 QY 241 MMSSTLRKADKELVHMTSAKKIPGVVELSLFDQYRLLESCEMVEYLMKGLMKRSIDHFGK 300
15 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
16 DB 249 MMSSTLRKADKELVHMTSAKKIPGVVELSLFDQYRLLESCEMVEYLMKGLMKRSIDHFGK 308
17 QY 301 LIFAPQVLDROEGKVEGILELFFMLLATTSRPELKIQHKRYLCVAMILLNSMPTL 360
18 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
19 DB 309 LIFAPQVLDROEGKVEGILELFFMLLATTSRPELKIQHKRYLCVAMILLNSMPTL 368
20 QY 361 VTATQDQSSPKLAHLINAVTDLVWYAKSGTSSQGSMPANILMLLSHVRH 414
21 IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII: IIILII:
22 DB 369 ATASQAEASSPKLTHIINAVTDLVWYAKSGTSSQGSMPANILMLLSHVRH 422

```

```

RESULT 9
US-08-764-870-12
1 Sequence 12; Application US/08/764,870
2 Patient No. 6236946
3 GENERAL INFORMATION:
4 APPLICANT: Scallan, Thomas S
5 APPLICANT: Baxter, John D
6 APPLICANT: Fletcher, Robert J
7 APPLICANT: Wagner, Richard L
8 APPLICANT: Kushner, Peter J
9 APPLICANT: Appleby, James W
10 APPLICANT: West, Brian
11 TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
12 NUMBER OF SEQUENCES: 16

```


385 VWIYAKSGTSSQGSMPANIIMLSHVRH 414
487 IHIMAKAGLTIDQOHRIADLIILSHIRH 516

RESULT 11

US-09-041-886-45
Sequence 45, Application US/09041886
Patent No. 6235872
GENERAL INFORMATION:
APPLICANT: Brodesen, Dale E.
APPLICANT: Radizadeh, Sharooz
TITLE OF INVENTION: Protophagic peptides, Dependent
TITLE OF INVENTION: Polypeptides and Methods of use
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09-041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 41,815
REFERENCE/DOCKET NUMBER: P-15 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-886-45

Query Match 47.84; Score 1054; DB 4; Length 595;

Best local Similarity 48.44; Pred. No. 9.7e-108;

Matches 218; Conservative 78; Mismatches 100; Indels 54; Gaps 10;

QY 10 TNEGGRGRO-----TSPVWLPPTGCHSP-LVHROLSHLYAEQ 50
DB 76 TGLPYGOSFAAANVNSMGLGCPPLNSVSPSLMLHPBPQLSPFLQPHGQVPTLNF 135
QY 51 KSPWC--FAASLEHTLVNRETLKRVKSNRCASPYIGD-----SKIDAHFCAVSDYA 103
DB 146 PSCTVTEAFAPPAVYP-----NSGNPQGCPEPLASTNCKSSMMESAKETPVAVNVA 192
QY 104 SGTIVYVWSCTPKAKFKRSTGSHNYTTPATWCTTEKRPKSTVACILSGYEVGMK 163
DB 193 SGHYVWSCTPKAKFKRSTGSHNYTTPATWCTTEKRPKSTVACILSGYEVGMK 252
QY 164 GGSPPRGCVPLVPSADFGCHCAKAPKPSGSHAPVP-----ELI 207
DB 253 GCTPKRPPGGMKIKKQPPH-----GPGGTFVNSATIMFAANIWSPSLMIPKSKNSIA 297
QY 208 DALSFEQIVLTLEAEFPHVLSR--ESAFTESMMMSI--TKIAKFTVHM--SWAKTAT 265
DB 408 LSLTADQMSVATIDAPP--IISYDPTPEFSAMMGLITNLAPETVHMIMNARVPD 366
QY 266 IVPLSTFDQVPLFSCVMEVIMMIMMPSIDHCKLIFATIVYTPDPRGVGGLPTD 325
DB 467 FVDLTHGVHLPFAMFPIIMGLVWSPSHVKKLPANLIIIPNCGTGVGAWP 426

QY 326 MLATTSREBELKCHREYCYAKAMLLNSSMYVIAL--QVANSKPIALINAVTAL 484
DB 427 MLATTSREPMNNIGREFFVTKSLTSSVYTESLTKSEKQIHVEVION 1101L 486
QY 385 VWIYAKSGTSSQGSMPANIIMLSHVRH 414
DB 487 IHIMAKAGLTIDQOHRIADLIILSHIRH 516

RESULT 12

US-08-453-998-2
Sequence 2, Application US/08453998
Patent No. 644438
GENERAL INFORMATION:
APPLICANT: CHAMON, PIERRE
APPLICANT: METZGER, DANIEL
APPLICANT: WHITE, JOHN
TITLE OF INVENTION: METHOD FOR THE PREPARATION OF A PROTEIN
TITLE OF INVENTION: BY YEASTS USING AN INDUCIBLE SYSTEM, VECTORS AND
TITLE OF INVENTION: CORRESPONDING TRANSFORMED STRAINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DABRY & CUSHMAN
STREET: 1100 NEW YORK AVE., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,998
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/161,064
FILING DATE: 03-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: CHAPIN, MARLANA K.
REGISTRATION NUMBER: 35,843
REFERENCE/DOCKET NUMBER: 1037/98493
TELEPHONE: 202-861-4711
TELEFAX: 202-822-0944
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-453-998-2

Query Match 47.84; Score 1054; DB 4; Length 595;

Best local Similarity 48.44; Pred. No. 9.7e-108;

Matches 218; Conservative 78; Mismatches 100; Indels 54; Gaps 10;

QY 10 TNEGGRGRO-----TSPVWLPPTGCHSP-LVHROLSHLYAEQ 50
DB 76 TGLPYGOSFAAANVNSMGLGCPPLNSVSPSLMLHPBPQLSPFLQPHGQVPTLNF 135
QY 51 KSPWC--FAASLEHTLVNRETLKRVKSNRCASPYIGD-----SKIDAHFCAVSDYA 103
DB 146 PSCTVTEAFAPPAVYP-----NSGNPQGCPEPLASTNCKSSMMESAKETPVAVNVA 192
QY 104 SGTIVYVWSCTPKAKFKRSTGSHNYTTPATWCTTEKRPKSTVACILSGYEVGMK 163
DB 193 SGHYVWSCTPKAKFKRSTGSHNYTTPATWCTTEKRPKSTVACILSGYEVGMK 252


```

US 09 249 645 1
Query Match 44 98% Score 967.5; DB 4; Length 229;
Post Local Similarity 94.78; Prod. No. 86-99;
Matches 197; Conservative 0; Mismatches 0; Indels 11; Gaps 2;

UY 208 DALSPQVLTLEAFEPHVLISRSAPTEFASMMSTELAKRELVIMISWAKKIDGCV 267
DB 1 DALSPQVLTLEAFEPHVLISRS ---ASMMSTELAKRELVIMISWAKKIDGCV 55

UY 268 ELSPFYGVLLSESMWYEMVLMKIMKESIDHPRKLLFADVLYDRPGKCVGELLEFDML 327
DB 56 ELSPFYGVLLSESMWYEMVLMKIMKESIDHPRKLLFADVLYDRPGKCVGELLEFDML 115

UY 328 LATSFRRELKLOHKREYLVKAMILLNNSMPLVATODADSRKLAHLINAVTALVWV 387
DB 116 LATSFRRELKLOHKREYLVKAMILLNNSMPLVATODADSRKLAHLINAVTALVWV 169

UY 408 IAKSGISSOOSMRLANLMLLSHVKA 415
DB 170 IAKSGISSOOSMRLANLMLLSHVKA 197

RESULT 19
US 08 846 620A 7
Sequence 7; Application US/088 66620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent to Release #1.0, Version #1.25 (EP9)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/846,620A
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP96/00943
FILING DATE:
APPLICATION NUMBER: 08-SEP-1995
FILING DATE: 08-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08-9605550.4
FILING DATE: 15-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08-9607532.0
FILING DATE: 11-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08-9609576.5
FILING DATE: 08-MAY-1996
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 226 amino acids
TYPE: amino acid
Topology: linear
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
US 08 846 620A 7

Query Match 45 78% Score 787; DB 2; Length 226;
Post Local Similarity 94.98; Prod. No. 7 66-79;
Matches 154; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

UY 252 ELVHIMSWAKKIDGVEVLSLPQVRLSESMWYEMVLMKIMKESIDHPRKLLFADVLYDR 311
DB 1 ELVHIMSWAKKIDGVEVLSLPQVRLSESMWYEMVLMKIMKESIDHPRKLLFADVLYDR 60

UY 312 DEKCVGELLEFDMLATSFRRELKLOHKREYLVKAMILLNNSMPLVATODADSR 371
DB 61 DEKCVGELLEFDMLATSFRRELKLOHKREYLVKAMILLNNSMPLVATODADSR 120

```

```

UY 472 KLAILINAVTALVWVIAKSGISSOOSMRLANLMLLSHVKA 414
DB 121 KLAILINAVTALVWVIAKSGISSOOSMRLANLMLLSHVKA 163

RESULT 19
US 08 564 264 1
Sequence 1; Application US/08564264
Patent No. 6040440
GENERAL INFORMATION:
APPLICANT: STEWART, Francis
TITLE OF INVENTION: REGULATION OF SITE-SPECIFIC
TITLE OF INVENTION: RECOMBINATION BY SITE-SPECIFIC RECOMBINASE/NUCLEAR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marumoto, Muray & Uram
STREET: 655 Fiftieth Street N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent to Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/564,264
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02088
FILING DATE: 28-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94 110 298.2
FILING DATE: 28-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Muray, Robert H.
REGISTRATION NUMBER: 22,980
REFERENCE/ID#-RET NUMBER: 4564 5019
TELEPHONE: (202)648-5000
TELEFAX: (202)648-4810
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 773 amino acids
TYPE: amino acid
Topology: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1-423)
OTHER INFORMATION: /note "FLP recombinase domain."
FEATURE:
NAME/KEY: misc_feature
LOCATION: (424-428)
OTHER INFORMATION: /note "Linker peptide."
FEATURE:
NAME/KEY: misc_feature
LOCATION: (429-773)
OTHER INFORMATION: /note "Estrogen binding domain."
US 08 564 264 1

Query Match 29 78% Score 695; DB 3; Length 773;
Post Local Similarity 44.48; Prod. No. 2 20-63;
Matches 143; Conservative 68; Mismatches 65; Indels 45; Gaps 9;

UY 121 KPSGSHINYICVATN-----QPTTK-----NPKRSQAQLPKQYEWGWRYSRRKQ 171
DB 491 KPSGSHINYICVATN-----QPTTK-----NPKRSQAQLPKQYEWGWRYSRRKQ 408

```


[illegible]

```

UY 184 FQULHAGKAKKSSGSHAPVR ELILDAISP FQVLTLLAEPRHV-LISRSAPETVAS 240
DB 218 -----NSLYNPDLVAPAKFPYKIVSHITVAETFTYMP-ETVDFSTIF 264
UY 241 MMMSLTKLAKREIVHISWAKKIPGVHSTFGVRLPSCWPELMIMMSIPDHK 400
DB 264 ALTIICDLAPRELVAITMAKHIPEFSTISLADQSLQSAAMELITLCVYRESLSEFE 424
UY 401 LIFAPDLVDRGKGVETFTFIMLIATISSEFELNIDKREYCVKAMILINS-SMYP 459
DB 424 IYVADYIMDFPWSK-LAVLITLNNATIGVKKYSMFLFEFEVTEKALIANSDSMH- 481
UY 460 IYVATQDADSSKRLAHIN-AVTPDALVWYIAKSGISSQDSQSRPLANILMLSHV 413
DB 482 -----IEDVAVOKLOLVIEHLDD-----YEAG-QIHEDPRAGKMDLIPLEK 425

RESULT 25
US-09-141-000-6
Sequence 6, Application US/09141000
Patent No. 6054295
GENERAL INFORMATION:
APPLICANT: Chen, Faoz
FILE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE OF INVENTION: RECEPTOR PROTEINS
FILE REFERENCE: 199909Y
CURRENT APPLICATION NUMBER: US/09/141,000
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 6
LENGTH: 418
TYPE: PRT
ORGANISM: Human
US-09-141-000-6

Query Match 26.0%; Score 572.5; DB 3; Length 418;
Best local Similarity 39.5%; Pred. No. 1,1e-54;
Matches 126; Conservative 70; Mismatches 90; Indels 32; Gaps 7.

UY 73 KKVSNRASPVTGSPSKRI-----AHFVAVSISYASQYAVSWSLDQKAPKRSQ 125
DB 98 RKLYDCSSSTIVDPQTKCFYMINSMKPELTVGDTLAGHYGVASCAWAKPKRTIQ 157
UY 126 GHNDYICPAPNOCITIKNRKKSQGNRLKCKCYEVCNVRGSR--PPGGYPLVPPQPSAD 183
DB 158 GNTVYSVPATNPFITIKPPKSSQAVPEMKCIKVMDLPGVPIQVPPQPPQYPPQPIAE 217
UY 184 FQULHAGKAKKSSGSHAPVR ELILDAISP FQVLTLLAEPRHV-LISRSAPETVAS 240
DB 218 -----NSLYNPDLVAPAKFPYKIVSHITVAETFTYMP-ETVDFSTIF 264
UY 241 MMMSLTKLAKREIVHISWAKKIPGVHSTFGVRLPSCWPELMIMMSIPDHK 400
DB 264 ALTIICDLAPRELVAITMAKHIPEFSTISLADQSLQSAAMELITLCVYRESLSEFE 424
UY 401 LIFAPDLVDRGKGVETFTFIMLIATISSEFELNIDKREYCVKAMILINS-SMYP 459
DB 424 IYVADYIMDFPWSK-LAVLITLNNATIGVKKYSMFLFEFEVTEKALIANSDSMH- 481
UY 460 IYVATQDADSSKRLAHIN-AVTPDALVWYIAKSGISSQDSQSRPLANILMLSHV 413
DB 482 -----IEDVAVOKLOLVIEHLDD-----YEAG-QIHEDPRAGKMDLIPLEK 425

RESULT 26
US-09-141-000-7
Sequence 2, Application US/09141000
Patent No. 6054295
GENERAL INFORMATION:
APPLICANT: Chen, Faoz
FILE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
FILE REFERENCE: 199909Y
CURRENT APPLICATION NUMBER: US/09/141,000
CURRENT FILING DATE: 1998-08-26
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 7
LENGTH: 500
TYPE: PRT
ORGANISM: Human
US-09-141-000-7

Query Match 26.0%; Score 572.5; DB 3; Length 500;
Best local Similarity 42.1%; Pred. No. 1.7e-54;
Matches 123; Conservative 56; Mismatches 81; Indels 32; Gaps 7.

UY 96 GAVSPYASVYHVGWSTGCKAKFKRSIGGNIYICPAINGCTIDKPPPSGACPLP 155
DB 103 CLVCGDIASVYHGVASGACACAKPKRTIQNMIEVSCPAINCHETFRKKKSCGACPKM 162
UY 156 -YHVS-NSR--PFR-YYDA VPPQPSAP-CLF-A-KAKNSD HADVPPHILITAI SFE 214
DB 163 CLKVGMKEVRLDPVPPQPSAPKPKPLDSESSPY-----STQISPPA 205
UY 214 GVLIT-----LIFAPRHV-LISRSAPETVASMMSTLADKELVHISWAKKIPGV 267
DB 206 KKLITKIVSLTVAPEFKIYAMP-PPQMPGADKALITICDIAPELIVITMAKHIHFS 265
UY 268 ELSTFWVPLISLQWPEVLMIMMSIPDHKRLHAFQVLPPEQDQVETLLEEDM 327
DB 266 SLNLSGQMSLQSAAMELITLCVYRESLSEFEVADYIMDFHSHR-LAGLELRYAI 324
UY 328 LATSSEPELKLKREYCVKAMILINS-SMYPITATQDADSSKRLAHIN 478
DB 325 LQIVRRYKKIKVKEFEVITKALIANSDSMV-----IEDLFAVOKLOLDLH 371

RESULT 27
US-08-816-620A-19
Sequence 17, Application US/0883620A
Patent No. 5958710
GENERAL INFORMATION:
APPLICANT:
FILE OF INVENTION: orphan receptor
NUMBER OF SEQUENCES: 19
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatCom in Relocase #1.0, Version #1.25 (EPP)
APPLICANT DATA:
FILING DATE: 08-MAY-1996
PCT/EP96/03934
PCT APPLICATION NUMBER:
PCT FILING DATE:
APPLICATION NUMBER: GR 9518272.1
FILING DATE: GR SEP-1995
PCT APPLICATION DATA:
PCT FILING DATE: 15-MAR-1996
APPLICATION NUMBER: GR 9605550.4
PCT APPLICATION DATA:
PCT FILING DATE: 11-APR-1996
APPLICATION NUMBER: GR 9605752.0
PCT APPLICATION DATA:
PCT FILING DATE: 08-MAY-1996
APPLICATION NUMBER: GR 9605756.5
PCT APPLICATION DATA:
PCT FILING DATE: 19-
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
TOPOLOGY: linear
ORIGINAL SOURCE:

```



```

UY 156 CTEVWVWVKSRRERGCVYKLVKROKSAFQVQWVWAKAKSKSGHAPVRELLDALSPGQ 214
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 61 CCGAVVAVLADIRKRRK - FKKVYVVALD----- AVALLPVGVGTINSGQR- -TSPSGQ 110
215 ---- LVLTLEAPPHVILISPSFA PFTAKMMRSTLKLAPLVIHMSWACEVGVV 247
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 111 LGLLPLPLINLMSIEFDVIVADINIKPLISNLSLISINOLGEBOLISVWWSKSLMRK 170
258 ELSLPLVQLRSCWVEMMCMIMMSLIDHQ- -KLIFAPQVILDRGKVEVHTLEFD 425
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 171 NLLHLLVLTLLVSSMSIWPFGIOWSVYKVSQGMVFAPDQIINFGPKK ESSGYSTL 229
426 MLATTSREPLKLGREYVGVKAMILNSSWYPIVATQDAVSSKLAHLINAVTALV 485
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 240 EFMOTLQEFVKLGVSSEFFLCKMKVLLINT- -PL- -FPLRSQSGFEHMSYTRRL 284
UY 486 WYIA - KSTISSQSGSKMLANL 406
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 284 KATLPRKGVSSSSQFVQVLRKL 407

RESULT 33
US 08 764 870-14
Sequence 14, Application US/08764870
Patient No. 6266946
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S.
APPLICANT: Baxter, John D.
APPLICANT: Fletcher, Robert J.
APPLICANT: Wagner, Richard L.
APPLICANT: Koshoff, Peter J.
APPLICANT: Apthoff, James W.
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
TITLE OF INVENTION: Binding Domains
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Coolidge Goodard
STREET: Five Palo Alto Square, 4000 El Camino Real
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94406
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc DOS/MS-DOS
SOFTWARE: Patient In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/764,870
FILING DATE: 13 Dec-1996
CLASSIFICATION: 540
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,540
FILING DATE: 13 Dec-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,544
FILING DATE: 13 Dec-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,606
FILING DATE: 14 Dec-1995
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Jackie N.
REGISTRATION NUMBER: 45,966
REFERENCE/Work Item Number: 07041, 246, 07106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)844,6000
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 943 amino acids
TYPE: amino acid
STRANDNESS:
TOPLOGY: linear

```

```

MULTIPLE TYPE: protein
US 08-764 870-14
Query Match 20.48; Score 448; 1b 4; Length 943;
Best Local Similarity 42.98; Prod. No. 2,50-40;
Matches 107; Conserved 63; Mismatches 127; Indels 28; Gaps 10;

UY 96 CAVCSYASGYGVGWSTVGRKAFKRSLSGHNQVGVATWQVTLGRKRSQACPIRK 155
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 567 QLGCPASDVRHGVLTQSTKVFRKADQJNLDAQNGCTVGRKRSQACPIRK 626

UY 156 CTEVWVWVKSRRERGCVYKLVKROKSAFQVQWVWAKAKSKSGHAPVRELLDALSPG 214
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 627 CCGAVVAVLADIRKRRK - FKKVYVVALDVAVALPVGLVINSQALSPG-----TSPG 678
214 Q- ---- LVLTLEAPPHVILISPSFA PFTAKMMRSTLKLAPLVIHMSWAKPIG 265
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 679 GQIGLTPPLINLMSIEFDVIVADINIKPLISNLSLISINOLGEBOLISVWWSKSLMR 748
266 EVELSLPQVQLRSCWVEMMCMIMMSLIDHQ- -KLIFAPQVILDRGKVEVHTLEF 424
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 749 FRLHLLPQVLTQVSSMSIWPFGIOWSVYKVSQGMVFAPDQIINFGPKK ESSGYSTL 797
424 FMLATTSREPLKLGREYVGVKAMILNSSWYPIVATQDAVSSKLAHLINAVT 481
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 798 EFMOTLQEFVKLGVSSEFFLCKMKVLLINTLPLEGRSQFEHMSYTRRL 855
UY 482 DALWVAVKSSQSGSKMLANL 406
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1b 856 - -IGLRKGVSSSSQFVQVLRKL 877

RESULT 34
US-08-980-115-14
Sequence 14, Application US/08980115
Patient No. 6266622
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S.
APPLICANT: Baxter, John D.
APPLICANT: Fletcher, Robert J.
APPLICANT: Wagner, Richard L.
APPLICANT: Koshoff, Peter J.
APPLICANT: Apthoff, James W.
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
TITLE OF INVENTION: 37AL 246/020US
CURRENT APPLICATION NUMBER: US/08/980,115
EARLIER FILING DATE: 1997-11-26
EARLIER APPLICATION NUMBER: 08/764,870
EARLIER FILING DATE: 1996-12-13
EARLIER APPLICATION NUMBER: 60/008,606
EARLIER FILING DATE: 1995-12-14
EARLIER APPLICATION NUMBER: 60/008,543
EARLIER FILING DATE: 1995-12-13
EARLIER APPLICATION NUMBER: 60/008,540
EARLIER FILING DATE: 1995-12-13
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patient In Ver. 2.0
SEQ ID NO: 14
LENGTH: 943
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME: 37AL 246/020US
LOCATION: (659) --(918)
OTHER INFORMATION: minimal ligand binding domain
US 08 980 115-14
Query Match 20.48; Score 448; 1b 4; Length 943;
Best Local Similarity 42.98; Prod. No. 2,50-40;
Matches 107; Conserved 63; Mismatches 127; Indels 28; Gaps 10;

```


LOCATION: (144)-(447)
 MOLECULE INFORMATION: minimal bound binding domain
 US-08 980 115, 16

Query Match 19,88; Score 437; Db 4; Length 452;

Best Local Similarity 28.6%; Prod. No. 1,60-49;

Matches 109; Conservative 74; Mismatches 107; Indels 92; Gaps 14;

DB 13 EGGKROT--TSRVLMP-----TGHLSPLVNHQLSHLYAEYKSNWEAKS -- 59

DB 20 GGLAGQSDPTAPVWVYGGVSKVP-YTSPYVKSEM-----GHMDSYSGPYGD 69

DB 60 -LEHTLPVNRRLTKRVSQNRVAPYVTSQSPVAPF-----CAVTSVVASGYH 108

DB 70 MLEETADHVLPT-----DYTPPQKTLTLDKESGCHY 104

DB 109 GWSGEGCKAFPKRSIGNHVYTPATNCTIUKNRKSCVGRLEPKYEVGVVVGKSKR 168

DB 105 GALTQGSQKVFPRKAAAGKOKYLAASKNCTIUKPRKNVPSRLKRYEAGMT-IGAR- 162

DB 169 EGGCYPLVPRQSADEPLHFAKAKRSQSHAPVREL-----LDALSPQLVTELEAF 224

DB 163 -----KIKKLGKIKLQERGFASSTSTSTETTKTKLYSHI EGYEGOPTLNVLEAF 214

DB 225 PIVLIS--KRSAPPTASMMKSTKLADEKLVIMISMAKIPGEVELSLFVRLLESVW 282

DB 215 PGVVCAGHDNNQDPSFALSLINELERQLVHVVAKALDGFRLHVDQMAVLOYSW 274

DB 283 MEVIMMLIMRSTIDHPRK--LIFAPLVLD-----RDRGKVECLDELPMILATISR 444

DB 275 MCLMVAFMCKSSFTNVSRLVFAPILVNFMHRSKRYSGVR-----MKHSQF 426

DB 444 FREKLQREKYLAVKAMILIN 454

DB 427 FCMQITTPQETLTKKALLLS 447

RESULT 47

US-09 041 886 11

Sequence 11, Application US/99041886

Patent No. 6,253,472

GENERAL INFORMATION:

APPLICANT: Hoesogen, Dale E.

APPLICANT: Redizadeh, Shantou

TITLE OF INVENTION: Protoprotic Peptides, Dependence

TITLE OF INVENTION: Peptideptides and Methods of Use

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell & Flores LLP

STREET: 4470 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM pc compatiblr

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/041,886

FILING DATE:

CLASSIFICATION:

AGENCY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 41,815

REFERENCE/PACKET NUMBER: P-147 2626

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 545-9601

TELEFAX: (619) 545-8949

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 918 amino acids

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09 041 886 11

Query Match 19,88; Score 437; Db 4; Length 918;

Best Local Similarity 28.6%; Prod. No. 40-39;

Matches 109; Conservative 74; Mismatches 107; Indels 92; Gaps 14;

DB 13 EGGKROT--TSRVLMP-----TGHLSPLVNHQLSHLYAEYKSNWEAKS -- 59

DB 486 GGLAGQSDPTAPVWVYGGVSKVP-YTSPYVKSEM-----GHMDSYSGPYGD 545

DB 60 -LEHTLPVNRRLTKRVSQNRVAPYVTSQSPVAPF-----CAVTSVVASGYH 108

DB 536 MLEETADHVLPT-----DYTPPQKTLTLDKESGCHY 570

DB 109 GWSGEGCKAFPKRSIGNHVYTPATNCTIUKNRKSCVGRLEPKYEVGVVVGKSKR 168

DB 571 GALTQGSQKVFPRKAAAGKOKYLAASKNCTIUKPRKNVPSRLKRYEAGMT-IGAR- 628

DB 169 EGGCYPLVPRQSADEPLHFAKAKRSQSHAPVREL-----LDALSPQLVTELEAF 224

DB 629 -----KIKKLGKIKLQERGFASSTSTSTETTKTKLYSHI EGYEGOPTLNVLEAF 680

DB 225 PIVLIS--KRSAPPTASMMKSTKLADEKLVIMISMAKIPGEVELSLFVRLLESVW 282

DB 681 PGVVCAGHDNNQDPSFALSLINELERQLVHVVAKALDGFRLHVDQMAVLOYSW 740

DB 283 MEVIMMLIMRSTIDHPRK--LIFAPLVLD-----RDRGKVECLDELPMILATISR 444

DB 741 MCLMVAFMCKSSFTNVSRLVFAPILVNFMHRSKRYSGVR-----MKHSQF 792

DB 444 FREKLQREKYLAVKAMILIN 454

DB 794 FCMQITTPQETLTKKALLLS 813

RESULT 48

US-08 764-870-15

Sequence 15, Application US/08764870

Patent No. 6,246,946

GENERAL INFORMATION:

APPLICANT: Scantland, Thomas S

APPLICANT: Baxter, John D

APPLICANT: Fletcher, Robert J

APPLICANT: Wagner, Richard L

APPLICANT: Kuchner, Peter J

APPLICANT: Arlotti, James W

TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand

TITLE OF INVENTION: Binding Domains

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Colony Goodward

STREET: Five Palo Alto Square, 4000 El Camino Real

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM pc compatiblr

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.40

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/764,870

FILING DATE: 13-DEC-1996

CLASSIFICATION: 540

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/008,540

FILING DATE: 13-DEC-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/008,543
 FILING DATE: 13-DEC-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/008,606
 FILING DATE: 14-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Nakamura, Jackie N
 REGISTRATION NUMBER: 35,966
 PREPARENCE/DOCKET NUMBER: USAL-246/0105
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650)843-5000
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 984 amino acids
 TYPE: amino acid
 STRANDBONDS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-764-870-15

Query Match: 19.5%; Score 430.5; DB 4; Length 984;
 Best Local Similarity 31.8%; Pred. No. 2,36-38;
 Matches 110; Conservative 59; Mismatches 140; Indels 37; Gaps 9;

40 PQLSHYARPKSPWFAFSLPHTLPVNP-----FTLKKVSGRCAFPVTPGSKRP 92
 542 QSTFQHLSSFTVWNILVSKSKSHQLSKRSQCYVLEKIPKVNSSLKRSVLT--GSSRP 599
 93 AHEGAVSYASVYHVNWSPTCAKPRSTVCHNDYTGATNCTTIDKAPRSCQAKR 152
 600 SKLTVQTPASCHGVAVTQSGYVFKPAVEGSHNYLCAPIINGLIDELPPKNTACR 659
 153 LKQTVNWKQSGHEKQGLVLEKQKASALDGLH-----AGKAKRSQGH 198
 660 LQKQIGADM-NIDAKRSK---KIQKIKGHEQDQDQPPPPPPSPSPGGITYIAAK 715
 199 APPVELLDALSPQVLT-----LLEAPPPVLTIS--PPSAPPTASMMSTIKLAK 251
 716 EFSVNTALVPOLSTSRALTSPVAVLENIPEIIVAGYDSKPTANLSTLNRLAK 775
 252 ELVHMTSMARKTIPGVEISSLFQVNLDSQWMEVLMGLMWSIDHCK--LTPADIVL 309
 776 QMIOVVKAKVLPGEKNIPLLEDQITLQYSWGLSSFAISWRSKYKHNISQFLYAPDLVF 835
 910 DRDEKVEGLEIFPMILATISPEPELKIQKEKLYKAMILNS 355
 836 N-EEKKHQSAMYEILQGMHQLSLQVFKQLITFEETTKVALLIST 880

RESULT 39
 US-08-980-115-15
 Sequence 16, Application US/08980115
 Patent No. 6266622
 GENERAL INFORMATION:
 APPLICANT: Scanlan, Thomas S.
 APPLICANT: Baxter, John D.
 APPLICANT: Fletchrick, Robert J.
 APPLICANT: Wagner, Richard L.
 APPLICANT: Kushner, Peter J.
 APPLICANT: Apiletti, James W.
 APPLICANT: West, Brian L.
 APPLICANT: Shulan, Andrew K.
 TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
 FILE REFERENCE: USAL-246/0205
 CURRENT APPLICATION NUMBER: US-08/980,115
 CURRENT FILING DATE: 1997-11-26
 EARLIER APPLICATION NUMBER: 08/764,870
 EARLIER FILING DATE: 1996-12-13
 EARLIER APPLICATION NUMBER: 60/008,606
 EARLIER FILING DATE: 1995-12-14
 EARLIER APPLICATION NUMBER: 60/008,543
 EARLIER FILING DATE: 1995-12-13
 EARLIER APPLICATION NUMBER: 60/008,540

EARLIER FILING DATE: 1995-12-13
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 15
 LENGTH: 984
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: DOMAIN
 LOCATION: (695)..(969)
 OTHER INFORMATION: minimal ligand binding domain
 US-08-980-115-15

Query Match: 19.5%; Score 430.5; DB 4; Length 984;
 Best Local Similarity 31.8%; Pred. No. 2,36-38;
 Matches 110; Conservative 59; Mismatches 140; Indels 37; Gaps 9;

40 PQLSHYARPKSPWFAFSLPHTLPVNP-----FTLKKVSGRCAFPVTPGSKRP 92
 542 QSTFQHLSSFTVWNILVSKSKSHQLSKRSQCYVLEKIPKVNSSLKRSVLT--GSSRP 599
 93 AHEGAVSYASVYHVNWSPTCAKPRSTVCHNDYTGATNCTTIDKAPRSCQAKR 152
 600 SKLTVQTPASCHGVAVTQSGYVFKPAVEGSHNYLCAPIINGLIDELPPKNTACR 659
 153 LKQTVNWKQSGHEKQGLVLEKQKASALDGLH-----AGKAKRSQGH 198
 660 LQKQIGADM-NIDAKRSK---KIQKIKGHEQDQDQPPPPPPSPSPGGITYIAAK 715
 199 APPVELLDALSPQVLT-----LLEAPPPVLTIS--PPSAPPTASMMSTIKLAK 251
 716 EFSVNTALVPOLSTSRALTSPVAVLENIPEIIVAGYDSKPTANLSTLNRLAK 775
 252 ELVHMTSMARKTIPGVEISSLFQVNLDSQWMEVLMGLMWSIDHCK--LTPADIVL 309
 776 QMIOVVKAKVLPGEKNIPLLEDQITLQYSWGLSSFAISWRSKYKHNISQFLYAPDLVF 835
 910 DRDEKVEGLEIFPMILATISPEPELKIQKEKLYKAMILNS 355
 836 N-EEKKHQSAMYEILQGMHQLSLQVFKQLITFEETTKVALLIST 880

RESULT 40
 US-08-764-870-13
 Sequence 13, Application US/08764870
 Patent No. 6236946
 GENERAL INFORMATION:
 APPLICANT: Scanlan, Thomas S.
 APPLICANT: Baxter, John D.
 APPLICANT: Fletchrick, Robert J.
 APPLICANT: Wagner, Richard L.
 APPLICANT: Kushner, Peter J.
 APPLICANT: Apiletti, James W.
 APPLICANT: West, Brian
 TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
 BINDING DOMAINS
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooley Goddard
 STREET: Five Palo Alto Square, 3000 El Camino Real
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306
 COMPUTER FILEABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICANT: Scanlan, Thomas S.
 FILING DATE: 13-DEC-1996
 CLASSIFICATION: 530

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/008,540
? FILING DATE: 18 DEC-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/008,544
? FILING DATE: 18 DEC-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/008,606
? FILING DATE: 14 DEC-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Nakamura, Jackie N
? REGISTRATION NUMBER: 55,766
? REFERENCE/WORK KEY NUMBER: 00/01-246/010S
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650)84-5000
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 777 amino acids
? TYPE: amino acid
? STRANDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US 08/764,870 14

```

```

Query Match 19,48; Score 426.5; DB 4; Length 777;
Host Local Similarity 41.68; Prod. No. 4,50-38;
Matches 114; Conserved 67; Mismatches 133; Indels 45; Gaps 12;

UY 73 KRYSNRVSP-----VGGSKRHAPVAVGSDYASGYHWASQECF 116
DB 486 KTVFSNRYSFMRGVSSPSSSATTTPPK-----GVSDPASYHWVITPSSF 441
UY 117 KAFKRSYCHNYITVATNOCTIDNRKRSPQACRLRKYEVGMVKGRRRRGYRLV 176
DB 442 KVFYKAVYQHNHYLFACHNDELIDKRRKNVACRYRKLQAZM-NLEAKRLKKLGL 500
UY 177 KRGRSADPOLRAGKRRKSGHAPVRELLDALSPQVLITLFAEPPIVLS--RPSA 244
DB 501 QQATGVSQ-----ETSEHGNKRTV PATIPQITP--TIVSLLEVLEPEVLYAGYDSSV 552
UY 245 PFEASMMSTETRLAKRELVIMISWKKIDCEVELSLPQVRLPSCWMEVIMMGLMWR 294
DB 554 POSTRLEMTLNMIDRGVYLAAYKAKALPGPRNHLIDQMTLQYSMMPLMAFALGRS 612
UY 295 L--DHWCKLFAHGLVLDDESKCVRGILLEPDMLLATTSRRRELQJHREYLVKAM 352
DB 613 YRGSANLILFAVDLILN FQRTIDPCWYDQKIMLYVSSELHIOVASYEYLVNKTLL 671
UY 354 LNSMYPVATVDALSSPKTAH--LNAVTVAVVATKSTSSQ--QDSMKLANLL 406
DB 672 LSS--VKNGLKSLQFLDILRMYINELAKALVKKGNSSQNMQKRYQNLIL 722

```

Search completed: June 16, 2003, 14:49:05
Job time: 21.4648 secs

Genome version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: Tue Jun 16, 2003, 12:15:18 : Search time 45.4127 seconds
(without alignments)
981 556 Million cells processed

Title: US-08-826-361a-6
Perfect score: 2203
Sequence: 1 MNVSPSNTNIPPTPTPT GQSMPTAMLMISVIVAP 416

Scoring table: BioSUM62
Gapop 10 0 Gapext 0 5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post processing: Maximum Match 08
Listing first 120 summaries

Database:

Published Applications-AA:
1. /seqn2_6/prodat/a/2/pub/paa/US08_NEW_PUB.pep.*
2. /seqn2_6/prodat/a/2/pub/paa/TCF_NEW_PUB.pep.*
3. /seqn2_6/prodat/a/2/pub/paa/US06_NEW_PUB.pep.*
4. /seqn2_6/prodat/a/2/pub/paa/US07_NEW_PUB.pep.*
5. /seqn2_6/prodat/a/2/pub/paa/US07_NEW_PUB.pep.*
6. /seqn2_6/prodat/a/2/pub/paa/US07_NEW_PUB.pep.*
7. /seqn2_6/prodat/a/2/pub/paa/US08_FUSCMB.pep.*
8. /seqn2_6/prodat/a/2/pub/paa/US08_FUSCMB.pep.*
9. /seqn2_6/prodat/a/2/pub/paa/US09_NEW_PUB.pep.*
10. /seqn2_6/prodat/a/2/pub/paa/US10_NEW_PUB.pep.*
11. /seqn2_6/prodat/a/2/pub/paa/US10_NEW_PUB.pep.*
12. /seqn2_6/prodat/a/2/pub/paa/US10_NEW_PUB.pep.*
13. /seqn2_6/prodat/a/2/pub/paa/US06_NEW_PUB.pep.*
14. /seqn2_6/prodat/a/2/pub/paa/US06_FUSCMB.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2198	99.8	530	9	US-10-198-785-2	Sequence 2, App11
2	1063	48.3	595	10	US-09-853-033-2	Sequence 2, App11
3	1058	48.0	701	9	US-10-052-092-12	Sequence 12, App1
4	1054	48.0	701	9	US-10-211-613-1	Sequence 1, App11
5	1056.5	48.0	575	10	US-09-893-666A-2	Sequence 2, App11
6	1054	47.8	595	9	US-10-096-710-1	Sequence 1, App11
7	1054	47.8	595	9	US-10-081-563-2	Sequence 2, App11
8	1054	47.8	595	9	US-10-052-092-9	Sequence 9, App11
9	1054	47.8	595	9	US-10-052-092-13	Sequence 13, App1
10	1054	47.8	595	9	US-10-052-092-14	Sequence 14, App1
11	1054	47.8	595	9	US-09-952-680A-10	Sequence 10, App1
12	1054	47.8	595	10	US-09-933-257A-2	Sequence 2, App11
13	1045	47.4	595	9	US-10-052-092-31	Sequence 31, App1
14	1044.5	47.4	595	9	US-10-052-092-40	Sequence 40, App1
15	1023	46.4	239	10	US-09-903-876-2	Sequence 2, App11
16	817.5	47.1	1099	9	US-10-259-864-4	Sequence 4, App11
17	779	45.4	1147	9	US-10-259-864-4	Sequence 4, App11
18	665.5	40.2	910	9	US-09-908-153B-40	Sequence 40, App1
19	665.5	30.2	922	9	US-09-608-153B-42	Sequence 42, App1

20	630	28.6	660	10	US-09-853-033-4	Sequence 4, App11
21	630	28.6	660	10	US-09-853-033-8	Sequence 8, App11
22	627	28.5	241	10	US-09-905-176-14	Sequence 14, App1
23	622	28.2	324	10	US-09-737-255-5	Sequence 5, App1
24	622	28.2	547	9	US-10-052-092-10	Sequence 10, App1
25	621	28.2	243	10	US-09-903-876-1	Sequence 1, App11
26	621	28.2	660	10	US-09-853-033-6	Sequence 6, App11
27	605.5	27.5	235	9	US-09-885-827-2	Sequence 2, App11
28	605.5	27.5	264	9	US-10-334-583-1	Sequence 1, App11
29	604.5	24.9	679	9	US-10-211-239-1	Sequence 1, App11
30	504	22.9	99	10	US-09-864-761-4873	Sequence 4873, A
31	457.5	20.8	689	9	US-10-302-845-2	Sequence 2, App11
32	426.5	19.4	727	9	US-10-153-668-226	Sequence 226, App
33	425	19.3	679	9	US-10-098-79A-2	Sequence 2, App11
34	423	19.2	1079	9	US-10-091-257-4	Sequence 2, App11
35	422.5	19.2	291	9	US-09-997-267-2	Sequence 2, App11
36	422.5	19.2	388	9	US-09-927-387-2	Sequence 2, App11
37	417.5	19.0	533	9	US-09-922-226-2	Sequence 2, App11
38	417.5	19.0	533	9	US-10-098-184-2	Sequence 2, App11
39	411.5	18.7	472	10	US-09-965-703-61	Sequence 61, App1
40	409	18.3	463	9	US-09-922-226-4	Sequence 4, App1
41	403	18.3	463	9	US-10-098-184-3	Sequence 3, App11
42	401.5	18.2	598	9	US-10-211-239-2	Sequence 2, App11
43	401.5	18.2	598	10	US-09-853-386-91	Sequence 91, App1
44	401.5	18.2	598	10	US-09-853-386-98	Sequence 98, App1
45	399.5	18.1	597	10	US-09-853-386-93	Sequence 93, App1
46	399	18.1	460	9	US-10-295-370-4	Sequence 4, App11
47	396.5	18.0	466	9	US-09-909-756-2	Sequence 2, App11
48	396.5	18.0	466	10	US-09-909-672-2	Sequence 2, App11
49	395	17.9	601	10	US-09-853-386-92	Sequence 92, App1
50	395	17.9	601	10	US-09-853-386-94	Sequence 94, App1
51	379.5	17.3	462	9	US-09-922-226-1	Sequence 1, App11
52	379.5	17.3	462	9	US-10-098-184-1	Sequence 1, App11
53	393	17.3	467	10	US-09-965-703-63	Sequence 63, App1
54	381.5	17.3	577	10	US-09-853-386-42	Sequence 42, App1
55	379	17.2	328	10	US-09-965-703-31	Sequence 31, App1
56	378	17.2	328	10	US-09-965-703-36	Sequence 36, App1
57	374	17.0	643	10	US-09-853-386-37	Sequence 37, App1
58	371	16.8	104	10	US-09-864-761-41519	Sequence 41519, A
59	67.4	16.7	592	10	US-09-853-386-34	Sequence 34, App1
60	66.5	16.6	628	10	US-09-853-386-66	Sequence 66, App1
61	66.5	16.6	628	10	US-09-853-386-68	Sequence 68, App1
62	365	16.5	628	10	US-09-853-386-73	Sequence 73, App1
63	365	16.5	628	10	US-10-005-165-4	Sequence 4, App11
64	364	16.5	598	10	US-09-853-386-37	Sequence 37, App1
65	364	16.5	598	10	US-09-853-386-38	Sequence 38, App1
66	364	16.5	598	10	US-09-853-386-39	Sequence 39, App1
67	364	16.5	598	10	US-09-853-386-41	Sequence 41, App1
68	363	16.5	626	10	US-09-853-386-64	Sequence 64, App1
69	362	16.4	598	10	US-09-853-386-40	Sequence 40, App1
70	361	16.4	598	10	US-09-853-386-32	Sequence 32, App1
71	361	16.4	598	10	US-09-853-386-35	Sequence 35, App1
72	361	16.4	598	10	US-09-853-386-36	Sequence 36, App1
73	361	16.4	598	10	US-09-853-386-45	Sequence 45, App1
74	360	16.3	626	10	US-09-853-386-65	Sequence 65, App1
75	360	16.3	626	10	US-09-853-386-96	Sequence 96, App1
76	359	16.3	627	10	US-09-853-386-67	Sequence 67, App1
77	359	16.2	627	10	US-10-005-169-2	Sequence 2, App11
78	358.5	16.3	461	10	US-09-731-557A-2	Sequence 2, App11
79	358	16.3	535	10	US-09-853-386-95	Sequence 95, App1
80	356.5	16.2	597	10	US-09-853-386-97	Sequence 97, App1
81	352.5	16.0	625	10	US-09-853-386-63	Sequence 63, App1
82	341.5	15.5	461	12	US-10-013-826-3	Sequence 3, App11
83	333.5	15.1	395	9	US-10-151-542A-21	Sequence 21, App1
84	333.5	15.1	454	9	US-09-814-604-3	Sequence 3, App11
85	333.5	15.1	454	10	US-09-797-727-2	Sequence 2, App11
86	333.5	15.1	462	9	US-09-814-604-1	Sequence 1, App11
87	333.5	15.1	462	10	US-09-797-727-3	Sequence 3, App11
88	333	15.1	61	9	US-09-814-604-21	Sequence 21, App1
89	332	15.1	61	9	US-09-922-226-7	Sequence 7, App11
90	333	15.1	61	9	US-09-922-226-169	Sequence 169, App
91	333	15.1	423	10	US-09-919-447-86	Sequence 86, App1
92	332	15.1	446	12	US-10-013-823-2	Sequence 2, App11

94 441 15.0 1089 9 US-10-259-864-2
 94 441 15.0 1090 9 US-10-259-864-6
 95 440 15.0 467 9 US-10-151-542A-2
 96 440 15.0 446 10 US-09-909-446-2
 97 440 15.0 446 10 US-09-909-425-2
 98 440 15.0 446 10 US-09-965-426-2
 99 428 14.9 513 10 US-09-965-703-5-9
 100 421.5 14.6 448 9 US-09-797-727-4
 101 421.5 14.6 746 9 US-09-042-488H-7
 102 420.5 14.5 625 10 US-09-965-703-16
 103 420.5 14.5 764 9 US-10-400-767-20
 104 418.5 14.5 448 9 US-09-814-604-2
 105 416 14.4 546 9 US-10-295-470-2
 106 412.5 14.2 746 9 US-10-300-757-18
 107 411.5 14.1 440 10 US-09-965-703-71
 108 408.5 14.0 746 9 US-09-042-488H-5
 109 407.5 14.0 1041 9 US-09-042-488H-9
 110 405.5 13.9 412 10 US-09-965-703-11
 111 405.5 13.9 412 10 US-09-965-703-12
 112 298 13.5 448 10 US-09-760-164-1
 113 295 13.4 476 9 US-10-188-721-1
 114 294 13.3 449 9 US-10-155-479-2
 115 292 13.3 359 10 US-09-760-164-9
 116 290 13.2 264 9 US-10-043-487-254
 117 290 13.2 496 10 US-09-965-703-3
 118 287.5 13.1 477 9 US-10-182-242-3
 119 277 12.6 458 10 US-09-760-164-2
 120 275 12.5 486 9 US-10-153-827-2

ALIGNMENTS

RESULT 1
 US-10-198-785-2
 Sequence 2: Application US/01098785
 Publication No. US2003002222A1

GENERAL INFORMATION:
 APPLICANT: OLYMPUS OPTICAL CO., LTD.
 TITLE OF INVENTION: Method of detecting binding reaction between protein and test sub-
 FILE REFERENCE: 7E1-0250538
 CURRENT APPLICATION NUMBER: US-10/198-785
 CURRENT FILING DATE: 2002-07-19
 PRIOR APPLICATION NUMBER: JP/2001-220444
 PRIOR FILING DATE: 2001-07-19
 PRIOR APPLICATION NUMBER: JP/2001-221063
 PRIOR FILING DATE: 2001-07-23
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: Patent In. Ver. 2.0
 SEQ ID NO: 2
 LENGTH: 510
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-198-785-2

Query Match 99.8% Score 2198; 106 G; Length 510;
 Local Similarity 100.0%; Prod. No. 2,1e-186;
 Matches 415; Conserved 0; Mismatches 0; Indels 0; Gaps 0

1 MNSLSNVNLEGGHGGELSPNVLWTPYHLSLWVHRLSHLYAHSKSPWFASS 60
 54 MNSLSNVNLEGGHGGELSPNVLWTPYHLSLWVHRLSHLYAHSKSPWFASS 114
 61 EHTLVNRETKKVSNGASVIVISKKIADHCVASVYASVYHGVNSVGGKAAE 120
 114 EHTLVNRETKKVSNGASVIVISKKIADHCVASVYASVYHGVNSVGGKAAE 174
 121 KPSLGHNIVTGANCTHHPNPPSSVATPLPGYVYAWGCVSSPPGCVLVPVGE 180
 174 KPSLGHNIVTGANCTHHPNPPSSVATPLPGYVYAWGCVSSPPGCVLVPVGE 234
 181 SAEGLHFAKKKSSAHAPVPELIDAISSPELVTLAAPPVVLISPSAPPTAS 240

18 234 SAEGLHFAKKKSSAHAPVPELIDAISSPELVTLAAPPVVLISPSAPPTAS 294
 241 MNSLSNVNLEGGHGGELSPNVLWTPYHLSLWVHRLSHLYAHSKSPWFASS 300
 294 MNSLSNVNLEGGHGGELSPNVLWTPYHLSLWVHRLSHLYAHSKSPWFASS 354
 301 LFAVDLVLDREKCVGLLEFDMLATSRREELKQREYCVKAMILLNSWYPL 360
 354 LFAVDLVLDREKCVGLLEFDMLATSRREELKQREYCVKAMILLNSWYPL 414
 361 VIATGVANSSSEKAILNNAVTAIVVAVAKSSVSSQSSMLANIMLSIVRIA 415
 414 VIATGVANSSSEKAILNNAVTAIVVAVAKSSVSSQSSMLANIMLSIVRIA 468

RESULT 2
 US-09-853-033-2
 Sequence 2: Application US/0985033
 Publication No. US2002010068A1
 GENERAL INFORMATION:
 APPLICANT: CHAMBER, PIERRE
 APPLICANT: METZGER, DANIEL
 TITLE OF INVENTION: TRANSGENIC MOUSE FOR TARGETED RECOMBINATION
 FILE REFERENCE: 06569170222
 CURRENT APPLICATION NUMBER: US/09/853-033
 CURRENT FILING DATE: 2001-05-11
 PRIOR APPLICATION NUMBER: FR-03/12570
 PRIOR FILING DATE: 2000-10-03
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: Patent In. Ver. 2.1
 SEQ ID NO: 2
 LENGTH: 595
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-853-033-2

Query Match 48.4% Score 1063; 106 G; Length 595;
 Local Similarity 48.7%; Prod. No. 9,1e-86;
 Matches 219; Conserved 78; Mismatches 99; Indels 54; Gaps 102

10 TNLGGHGGH-----TSPNVLWTPYHLSLWVHRLSHLYAHSKSPWFASS 50
 76 TNLGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGH 135
 91 KSNV---EASLEHTLVNRETKKVSNGASVIVISKKIADHCVASVYASVYHGVNSVGGKAAE 104
 136 TNLGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGH 192
 104 SGNVGVNSGEGKAKFKPSVGGHNDYVGVNGLLNKRRSGVGRRLKRYGVGVK 164
 193 SGNVGVNSGEGKAKFKPSVGGHNDYVGVNGLLNKRRSGVGRRLKRYGVGVK 252
 164 GGRREKCVGLVRRKSDLEGLFAKKKSSAHAPVPELIDAISSPELVTLAAPPVVLISPSAPPTAS 207
 254 GGRREKCVGLVRRKSDLEGLFAKKKSSAHAPVPELIDAISSPELVTLAAPPVVLISPSAPPTAS 307
 208 DATSPGLVTLAAPPVVLISPSAPPTASMSSTKRLAKELVHMTSAKATPG 265
 408 DATSPGLVTLAAPPVVLISPSAPPTASMSSTKRLAKELVHMTSAKATPG 466
 266 EYLSLFDVRLLESVHMTVLMGMMMSIDHPRKLFAPVIVLDREKCVGLLEFD 325
 467 EYLSLFDVRLLESVHMTVLMGMMMSIDHPRKLFAPVIVLDREKCVGLLEFD 426
 429 MLALISPEELIKREYCVKAMILLNSWYPLVAVVAVAKSSVSSQSSMLANIMLSIVRIA 484
 427 MLALISPEELIKREYCVKAMILLNSWYPLVAVVAVAKSSVSSQSSMLANIMLSIVRIA 486
 485 VVAVAKSSVSSQSSMLANIMLSIVRIA 414
 487 INMAKAVTLDGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGHGGH 516


```
QY 426 MLATSPRPETKIGKPYGVKAMIIINSSMPIVTATGQAGSGPATANHINVTAT 194
      |||||:||||:|:|||||:|||||:|:||||:|:||||:|:||||:|:||||:|
Db 427 MLATSPRPMMNIQSPVPTKSTIIINSVTLSTKSLKEEMHHPVDEKTIYTL 186
      :|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|
QY 485 VWVIAKSGISSQOOSMRLANLMLLSHVRH 414
      :|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|
Db 487 IILMAKAGITLQOOHQRLAQLLILSHIRH 516
      :|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|

RESULT 8
US-10-052-092-9
: Sequence 9, Application US/10052092
: Publication No. US2004092778A1
: GENERAL INFORMATION:
: APPLICANT: Fujita, Suzanne
: APPLICANT: Alfred, D.
: APPLICANT: Hopp, Torsten A.
: APPLICANT: O'Connell, Peter
: TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapies
: FILE REFERENCE: P02102052
: CURRENT APPLICATION NUMBER: US/10/052,092
: PRIOR FILING DATE: 2002-01-18
: PRIOR APPLICATION NUMBER: US 60/262,990
: PRIOR FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: US 60/304,018
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: Patent In version 3.1
: SEQ ID NO 9
: LENGTH: 595
: TYPE: PRI
: ORGANISM: Human
US-10-052-092-9

Query Match 47.88; Score 1054; DB 9; Length 595;
Best Local Similarity 48.48; Pred No 5, 7e-85;
Matches 218; Conservative 78; Mismatches 100; Indels 54; Gaps 10;

QY 10 TNEGGPGKQ-----TSPVLMPTQCHLSP-LVHROLSHLYARQ 50
      |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 76 TGLPYGQSEAAAFGNSGLGPPPLNSVSPSLMLLHPPOLSPTLQPGQGVYTLNE 135
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 51 KSPWC--FAKSLFHTLPVNETLKPKVSNRCASPYTGQ-----SKRAHCAVCSYA 103
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 136 PSQTVARPAAPPAFYR---NSINRGSGREKLASTINDGSMAMESAKETRYCAVCNOYA 192
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 104 SGYHVYWSCEGCAKAFKRSIGCHNYICTAINCTTIANPKRSQAVTLEKCYEVGMVA 163
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 193 SGYHVYWSCEGCAKAFKRSIGCHNYICTAINCTTIANPKRSQAVTLEKCYEVGMVA 252
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
QY 164 GGSRRPGCYRIYVRGSADEQLHCAGKAKRSCHAPRY-----PIL 207
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 253 GGTIRDRPRGPMKIKRQGRD-----GEJGEVGSAGDMVANIWDSPIMRKSKNSTA 307
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 208 DALSPDVLITLLEPHVILISR--PSAPPTASMMMSLTKIAKELVHMSMAKRIHG 265
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 308 LSLTADQVNSALIDAPP-ILYSEYDPTREPSASMMGLITINADRELVHMINMAKRVG 366
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 266 FVELSTFDVYLLLESTWMEVLMKIMKRSIDHGKLIAPADLVITPEDEKCYEGHLEFD 325
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 367 FVDLTLDQVHLLEAMLEIMGLVWRSMEHVKLIAPANILIDRNOCKCYEGHLEFD 426
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
QY 326 MLATSPRPETKIGKPYGVKAMIIINSSMPIVTATGQAGSGPATANHINVTAT 284
      |||||:||||:|:|||||:|||||:|:||||:|:||||:|:||||:|:||||:|
Db 427 MLATSPRPMMNIQSPVPTKSTIIINSVTLSTKSLKEEMHHPVDEKTIYTL 486
      :|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|
QY 485 VWVIAKSGISSQOOSMRLANLMLLSHVRH 414
      :|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|
Db 487 IILMAKAGITLQOOHQRLAQLLILSHIRH 516
      :|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|||||:|:|:|:|

RESULT 9
```

```
US-10-052-092-13
: Sequence 13, Application US/10052092
: Publication No. US2004092778A1
: GENERAL INFORMATION:
: APPLICANT: Fujita, Suzanne
: APPLICANT: Alfred, D.
: APPLICANT: Hopp, Torsten A.
: APPLICANT: O'Connell, Peter
: TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapies
: FILE REFERENCE: P02102052
: CURRENT APPLICATION NUMBER: US/10/052,092
: PRIOR FILING DATE: 2002-01-18
: PRIOR APPLICATION NUMBER: US 60/262,990
: PRIOR FILING DATE: 2001-01-19
: PRIOR APPLICATION NUMBER: US 60/304,018
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: Patent In version 3.1
: SEQ ID NO 13
: LENGTH: 595
: TYPE: PRI
: ORGANISM: human
US-10-052-092-13

Query Match 47.88; Score 1054; DB 9; Length 595;
Best Local Similarity 48.48; Pred No 5, 7e-85;
Matches 218; Conservative 78; Mismatches 100; Indels 54; Gaps 10;

QY 10 TNEGGPGKQ-----TSPVLMPTQCHLSP-LVHROLSHLYARQ 50
      |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 76 TGLPYGQSEAAAFGNSGLGPPPLNSVSPSLMLLHPPOLSPTLQPGQGVYTLNE 135
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
QY 51 KSPWC--FAKSLFHTLPVNETLKPKVSNRCASPYTGQ-----SKRAHCAVCSYA 103
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 136 PSQTVARPAAPPAFYR---NSINRGSGREKLASTINDGSMAMESAKETRYCAVCNOYA 192
      :|:|:|:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 104 SGYHVYWSCEGCAKAFKRSIGCHNYICTAINCTTIANPKRSQAVTLEKCYEVGMVA 163
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 193 SGYHVYWSCEGCAKAFKRSIGCHNYICTAINCTTIANPKRSQAVTLEKCYEVGMVA 252
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
QY 164 GGSRRPGCYRIYVRGSADEQLHCAGKAKRSCHAPRY-----PIL 207
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 253 GGTIRDRPRGPMKIKRQGRD-----GEJGEVGSAGDMVANIWDSPIMRKSKNSTA 307
      :|||:|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 208 DALSPDVLITLLEPHVILISR--PSAPPTASMMMSLTKIAKELVHMSMAKRIHG 265
      :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 308 LSLTADQVNSALIDAPP-ILYSEYDPTREPSASMMGLITINADRELVHMINMAKRVG 366
      :|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 266 FVELSTFDVYLLLESTWMEVLMKIMKRSIDHGKLIAPADLVITPEDEKCYEGHLEFD 325
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
Db 367 FVDLTLDQVHLLEAMLEIMGLVWRSMEHVKLIAPANILIDRNOCKCYEGHLEFD 426
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
QY 326 MLATSPRPETKIGKPYGVKAMIIINSSMPIVTATGQAGSGPATANHINVTAT 484
      |||||:||||:|:|||||:|||||:|:||||:~|:||||:~|:||||:~|:||||:~|
Db 427 MLATSPRPMMNIQSPVPTKSTIIINSVTLSTKSLKEEMHHPVDEKTIYTL 486
      :|||||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
QY 485 VWVIAKSGISSQOOSMRLANLMLLSHVRH 414
      :|||||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|
Db 487 IILMAKAGITLQOOHQRLAQLLILSHIRH 516
      :|||||:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|:~|

RESULT 10
US-10-052-092-14
: Sequence 14, Application US/10052092
: Publication No. US2004092778A1
: GENERAL INFORMATION:
: APPLICANT: Fujita, Suzanne
: APPLICANT: Alfred, D.
: APPLICANT: Hopp, Torsten A.
: APPLICANT: O'Connell, Peter
: TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapies
: FILE REFERENCE: P02102052
: CURRENT APPLICATION NUMBER: US/10/052,092
```


Db 76 TGLPYGSGEAAFGSGNGLGPPPLNSVSPRLMLHPPOI:SPRIQHGQGVVPEYLE 135
UY 51 KSIWGC--EARSLEHTLPVNRKELIKRVSQNGCASPYTGQ-----SKRUARPCAVCSYA 103
Db 136 PSQVYFAEPAPAFVPP-----NSDNPRGGREPLASINDGSMAMSAKETPCVAGCINUYA 192
UY 104 SGHYGVWVSGCKAKFKRSIGCHNDYTPATNCGTLENPKRSVQACHLKYCYEVGMK 163
Db 193 SGHYGVWVSGCKAKFKRSIGCHNDYTPATNCGTLENPKRSVQACHLKYCYEVGMK 252
UY 164 CGSFRPGCYRIVPPPSADPQHFGAKAPSSGAPVPP-----EELI 207
Db 253 CGIPKIPRGYRMKIKRPPDP-----CGPPGCVGSGCTMPRANINSPRI:MPYKPKYLA 207
UY 208 DALSPQOLVITLFAERPVHLSR--PSAPFTASMMMSITKLAKEIVHMSMARKITG 265
Db 408 LSTADQMSALLDAEHP--LTVSEYDTPRFSASMMGLTINLADRELVHMI:NMARVHG 366
UY 266 FVETLSLEPDVQLLESCWMEVIMKIMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 325
Db 467 FVETLSLEPDVQLLESCWMEVIMKIMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 426
UY 426 MLATTSPPPEIKAKRKYIVVAMTINSSMYRTVAT--QVANSSEPAITNAVTAT 394
Db 427 MLATTSPPPEIKAKRKYIVVAMTINSSMYRTVAT--QVANSSEPAITNAVTAT 464
UY 485 VWIARSGISGQOOSMRILANLMLSHVRH 414
Db 487 IHMAKAGLTIDOOHRIKRLAQLLLLSHVRH 516

RESULT 13
US-10-052-092-31
Sequence 32, Application: US 7/10/02 2092
Publication No. US2003002778A1
GENERAL INFORMATION:
APPLICANT: Futaba, Suzanne
APPLICANT: Allred, D.
APPLICANT: Hopp, Torsten A.
APPLICANT: O'Connell, Peter
TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapy
FILE REFERENCE: P02102052
CURRENT APPLICATION NUMBER: US 7/10/02 2092
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/262,990
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/304,018
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patent In version 3.1
SEQ ID NO 31
LENGTH: 595
TYPE: PRT
ORGANISM: pig
US-10-052-092-31

Query Match 47.4%: Score 1045; DB 9; Length 595;
Best Local Similarity 48.0%: Pred. No. 3,60-84;
Matches 217; Conservative 80; Mismatches 97; Indels 56; Gaps 12;

UY 10 TLEGGPGKQ-----TTPSWVWIPJGHLSP-LVHROLSHU:YAEPU 50
Db 76 SGLAYGPGSAAFGANGGLGPPPLNSVSPRLMLHPPOI:SPRIQHGQGVVPEYLE 135
UY 51 KSIWGC--EARSLEHTLPVNRKELIKRVSQNGCASPYTGQ-----SKRUARPCAVCSYA 103
Db 146 PSQVYFAEPAPAFVPP-----NSDNPRGGREPLASINDGSMAMSAKETPCVAGCINUYA 192
UY 104 SGHYGVWVSGCKAKFKRSIGCHNDYTPATNCGTLENPKRSVQACHLKYCYEVGMK 163
Db 193 SGHYGVWVSGCKAKFKRSIGCHNDYTPATNCGTLENPKRSVQACHLKYCYEVGMK 252
UY 164 CGSFRPGCYRIVPPPSADPQHFGAKAPSSGAPVPP-----EELI 207
Db 253 CGIPKIPRGYRMKIKRPPDP-----CGPPGCVGSGCTMPRANINSPRI:MPYKPKYLA 207
UY 208 DALSPQOLVITLFAERPVHLSR--PSAPFTASMMMSITKLAKEIVHMSMARKITG 265
Db 408 LSTADQMSALLDAEHP--LTVSEYDTPRFSASMMGLTINLADRELVHMI:NMARVHG 366
UY 266 FVETLSLEPDVQLLESCWMEVIMKIMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 325
Db 467 FVETLSLEPDVQLLESCWMEVIMKIMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 426
UY 426 MLATTSPPPEIKAKRKYIVVAMTINSSMYRTVAT--QVANSSEPAITNAVTAT 394
Db 427 MLATTSPPPEIKAKRKYIVVAMTINSSMYRTVAT--QVANSSEPAITNAVTAT 464
UY 485 VWIARSGISGQOOSMRILANLMLSHVRH 414
Db 487 IHMAKAGLTIDOOHRIKRLAQLLLLSHVRH 516

Db 253 GGIIPKIPRGYRMKIKRPPDP-----CGPPGCVGSGCTMPRANINSPRI:MPYKPKYLA 207
UY 208 DALSPQOLVITLFAERPVHLSR--PSAPFTASMMMSITKLAKEIVHMSMARKITG 265
Db 408 LSTADQMSALLDAEHP--LTVSEYDTPRFSASMMGLTINLADRELVHMI:NMARVHG 366
UY 266 FVETLSLEPDVQLLESCWMEVIMKIMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 325
Db 467 FVETLSLEPDVQLLESCWMEVIMKIMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 426
UY 426 MLATTSPPPEIKAKRKYIVVAMTINSSMYRTVAT--QVANSSEPAITNAVTAT 394
Db 427 MLATTSPPPEIKAKRKYIVVAMTINSSMYRTVAT--QVANSSEPAITNAVTAT 464
UY 485 VWIARSGISGQOOSMRILANLMLSHVRH 414
Db 487 IHMAKAGLTIDOOHRIKRLAQLLLLSHVRH 516

RESULT 14
US-10-052-092-30
Sequence 39, Application: US 7/10/02 2092
Publication No. US2003002778A1
GENERAL INFORMATION:
APPLICANT: Futaba, Suzanne
APPLICANT: Allred, D.
APPLICANT: Hopp, Torsten A.
APPLICANT: O'Connell, Peter
TITLE OF INVENTION: Methods and Composition in Breast Cancer Diagnosis and Therapy
FILE REFERENCE: P02102052
CURRENT APPLICATION NUMBER: US 7/10/02 2092
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: US 60/262,990
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/304,018
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patent In version 3.1
SEQ ID NO 30
LENGTH: 595
TYPE: PRT
ORGANISM: hamster
US-10-052-092-30

Query Match 47.4%: Score 1044.5; DB 9; Length 595;
Best Local Similarity 49.9%: Pred. No. 40-84;
Matches 211; Conservative 80; Mismatches 105; Indels 27; Gaps 9;

UY 16 PGROTSPN--VIMPTGHLSP-LVHROLSHU:YAEPU:KSIWGC--EARSLEHTLPVNRKELIKRVSQNGCASPYTGQ-----SKRUARPCAVCSYA 126
Db 97 POLNSVSPRLMLHPPOI:SPRIQHGQGVVPEYLE:PSQVYFAEPAPAFVPP-----NSDNPRGGREPLASINDGSMAMSAKETPCVAGCINUYA 215
UY 72 KRYVSNPFAVPTGDSK-----PQAHPCAVSDVASYVHYVWVSGCKAKFKRSIG 126
Db 156 NPSQVFAEPAPAFVPP-----NSDNPRGGREPLASINDGSMAMSAKETPCVAGCINUYA 215
UY 127 HNTYTPATNCGTLENPKRSVQACHLKYCYEVGMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 325
Db 216 HNTYTPATNCGTLENPKRSVQACHLKYCYEVGMKRSIDHGRKLEAPDVLPDRKCYEGLLEFD 426
UY 185 --GLHFAKAKRS-----GHAAPVDELIMLSPQOLVITLFAERPVHLSR 242
Db 276 GRNDQMSALLDAEHP--LTVSEYDTPRFSASMMGLTINLADRELVHMI:NMARVHG 366
UY 233 SAFTASMMMSITKLAKEIVHMSMARKITG 265
Db 434 SRFPSASMMMSITKLAKEIVHMSMARKITG 265
UY 293 RSIHGRKLEAPDVLPDRKCYEGLLEFD 325
Db 494 RSIHGRKLEAPDVLPDRKCYEGLLEFD 426

Db 789 TIVPAALPOLTP--TIVNLTIVTPPEVIYACVNSSSPGPPITISFPAHPTGPPASMM 846
 QY 214 GTPV 1A0PT 27
 Db 847 LTNEPEVLYAGVSSVSSKGLSSKAHOPFPAICQIAKYTNNSGPOVSDIHLADRE 906
 QY 254 VHMISMAKIPGVEVLSIFQVLLFSGVMEVIMMPSIDHCKLIFAPGVLIDRE 314
 Db 907 VHMINKAKVPPCHVLLTIDVLLFCAMLETIMGLVMSRHHVLLFAPNILLIDRE 966
 QY 314 GQVETLEFEMILATTSPPETKIGHEVLYVAMILNNSMYPLVAT--GDASSPK 372
 Db 967 GQVETLEFEMILATTSPPETKIGHEVLYVAMILNNSMYPLVAT--GDASSPK 372
 QY 373 LAILLNAVTDALVWIAKSGISSQVSMKLANIMLISHVRH 414
 Db 1027 IIRVITIKITDILIMMAKAGILGCHQRLAJALLISHKH 1098

RESULT 18
 US-09-908-153B-40
 ? Sequence 40, Application US/09/908153B
 ? Patent No. US2002016871A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Barbas, Carlos F.
 ? APPLICANT: Heerli, Roger
 ? APPLICANT: Schopfer, Ulrich
 ? TITLE OF INVENTION: REGULATION OF GENE EXPRESSION USING
 ? TITLE OF INVENTION: SINGLE-CHAIN, MONOMERIC, LIGAND DEPENDENT POLYPEPTIDE
 ? FILE REFERENCE: ISRI 725.1
 ? CURRENT APPLICATION NUMBER: US/09/908,153B
 ? CURRENT FILING DATE: 2001-07-18
 ? PRIOR APPLICATION NUMBER: US 09/619,063
 ? PRIOR FILING DATE: 2000-07-18
 ? NUMBER OF SEQ ID NOS: 54
 ? SOFTWARE: FastSeq for Windows Version 4.0
 ? SEQ ID NO 40
 ? LENGTH: 910
 ? TYPE: PRT
 ? ORGANISM: Unknown
 ? FEATURE:
 ? OTHER INFORMATION: Synthesized
 US-09-908-153B-40

Query Match 30.2% Score 665.5, Db 9, Length 910
 Host Local Similarity 38.1% Pred. No. 2,7e-50
 Matches 164, Conservative 76, Mismatches 106, Indels 85, Gaps 17
 QY 47 VVHPOUSHYAPDQKSTWTFEASTFHPIPVNFETIKPVVSNPQVAFVCTGSEAFAP 96
 Db 42 LVHGFTHGFEKYPVPTGKSPGQSSSHVHODPTGGERPYCP--EY----- 134
 QY 97 AVCSYVA--SGYIVG--VMSCEG--KAF-----PKSTGHNHYTCFATNQITL 144
 Db 83 SDPKLAPHQPTHTGKPYACAPPCGKSPGQSSSHVHODPTGGERPYCP--EY----- 134
 QY 144 PPKSGACR-----LPKCYEVGMYKQSPPEPGYPLVPRQSGADPLHVG 190
 Db 145 -CKSPSTGPIAHGHTHTGKPYCPFGG--FSTSPG-----KLVGQPT----- 179
 QY 191 KAKRSGHAPRVEFLID-----ALSPQVITTFAPDP 225
 Db 180 TPKRTSGQAPRPMACVMEAVNIWISPLVITHTFNSIAISLIDGMSVALIDAP 229
 QY 226 HVLIS-PPSAPTEASMMMSLTKIDKRELVHMSMAKRIQGVPSIFPVVLLSSQME 284
 Db 240 MYSFVPPSPPESEASMMGLITNLADPELVHMINNAKRPVSGDLENHIDVHILPCAM 299
 QY 295 VIMMELMWSIDHCKLIFAPDVLIDPQKCVFGLLEFIMLATSFFELKQHEV 344
 Db 400 LLMGLVMSRHHVLLFAPNILLIDRNGKCVGVNVEIFDMILATSSFRMNLQGEFF 359

QY 445 LCVKAMILNNSMYPLVAT--GDASSPKLAILNAVTDALVWIAKSGISSQVSMK 404
 Db 404 NIMLISHVRH 414
 QY 420 QILLILSHRH 430

RESULT 19
 US-09-908-153B-42
 ? Sequence 42, Application US/09/908153B
 ? Patent No. US2002016871A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Barbas, Carlos F.
 ? APPLICANT: Heerli, Roger
 ? APPLICANT: Schopfer, Ulrich
 ? TITLE OF INVENTION: REGULATION OF GENE EXPRESSION USING
 ? TITLE OF INVENTION: SINGLE-CHAIN, MONOMERIC, LIGAND DEPENDENT POLYPEPTIDE
 ? FILE REFERENCE: ISRI 725.1
 ? CURRENT APPLICATION NUMBER: US/09/908,153B
 ? CURRENT FILING DATE: 2001-07-18
 ? PRIOR APPLICATION NUMBER: US 09/619,063
 ? PRIOR FILING DATE: 2000-07-18
 ? NUMBER OF SEQ ID NOS: 53
 ? SOFTWARE: FastSeq for Windows Version 4.0
 ? SEQ ID NO 42
 ? LENGTH: 922
 ? TYPE: PRT
 ? ORGANISM: Unknown
 ? FEATURE:
 ? OTHER INFORMATION: Synthesized
 US-09-908-153B-42

Query Match 30.2% Score 665.5, Db 9, Length 922
 Host Local Similarity 38.1% Pred. No. 2,8e-50
 Matches 164, Conservative 76, Mismatches 106, Indels 85, Gaps 17
 QY 47 VVHPOUSHYAPDQKSTWTFEASTFHPIPVNFETIKPVVSNPQVAFVCTGSEAFAP 96
 Db 32 LVHGFTHGFEKYPVPTGKSPGQSSSHVHODPTGGERPYCP--EY----- 134
 QY 97 AVCSYVA--SGYIVG--VMSCEG--KAF-----PKSTGHNHYTCFATNQITL 144
 Db 83 SDPKLAPHQPTHTGKPYACAPPCGKSPGQSSSHVHODPTGGERPYCP--EY----- 134
 QY 144 PPKSGACR-----LPKCYEVGMYKQSPPEPGYPLVPRQSGADPLHVG 190
 Db 135 -CKSPSTGPIAHGHTHTGKPYCPFGG--FSTSPG-----KLVGQPT----- 179
 QY 191 KAKRSGHAPRVEFLID-----ALSPQVITTFAPDP 225
 Db 180 TPKRTSGQAPRPMACVMEAVNIWISPLVITHTFNSIAISLIDGMSVALIDAP 229
 QY 226 HVLIS-PPSAPTEASMMMSLTKIDKRELVHMSMAKRIQGVPSIFPVVLLSSQME 284
 Db 240 MYSFVPPSPPESEASMMGLITNLADPELVHMINNAKRPVSGDLENHIDVHILPCAM 299
 QY 295 VIMMELMWSIDHCKLIFAPDVLIDPQKCVFGLLEFIMLATSFFELKQHEV 344
 Db 445 LCVKAMILNNSMYPLVAT--GDASSPKLAILNAVTDALVWIAKSGISSQVSMK 404
 QY 404 NIMLISHVRH 414
 Db 420 QILLILSHRH 430

RESULT 20


```

Hager Ph.D., Gordon L.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MONITORING
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Neelie S. Rosenberg
: STREET: 127 Peachtree Street, Suite 1200
: CITY: Atlanta
: STATE: Georgia
: COUNTRY: USA
: ZIP: 30303
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/091,406
: FILING DATE: 15-Nov-02 US20020165381A1
: CLASSIFICATION: -Unknown-
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/908,373
: FILING DATE: 08-Dec-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Spiby, Elizabeth
: REGISTRATION NUMBER: 38298
: REFERENCE TO PRIOR ART:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 404-668-0770
: TELEFAX: 404-668-9880
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1070 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-001-4868-2

Query Match 19.2% Score 422.5; Db 9; Length 1070;
Best Local Similarity 32.2%; Pred. No. 1,16 28;
Matches 107; Conservative 63; Mismatches 136; Indels 26; Gaps 9;

QY 81 ASPTGDSKRAHFAVAGSYVAGYHGWSCGKAPFKRSIGGNDYGPATNCTI 140
DB 704 SSATGTPPK-----LGLVSDIFASCTHGVITGSKVFKKAVDQGNHILASGNDCT 11
141 PNPKPSQA*PLFK*YEV*MKV*TSFEFF*YPLVPEQSAL*E*Q*W*AK*VS*Q*MAI* 259
1760 DKIPKNC*PACPKYK*IG*W*-----NI*FAKTKKK*IG*IG*Q*ATG*G*V*Q*TS*P*ND* 911
QY 201 PVFPIILDAISPPQVITTEAFPPHVIS--EPKAPTEASMMMSITKLAK*ELVMS 259
DB 812 TIVPAALPQITP--TIVSILFVTEPVIYAVYDSSVPSVAMPIMTITNM*G*P*OV*IA*V* 869
259 WAKKIPGEVELSTFQVRLSCGMMELMGLMMSI-DHDKLI-FAFVLEBDEGNC 316
870 WAAALITGIPNHIHQMTLLGYSWPFIAPALQWSPYSSNLTGFAPELLIN*Q*MS 929
QY 317 VGTITFIPMLAITSFPEELKIQKKEIVCYKAMILLNSMKPLVATQVDSPELAH 476
DB 929 LKGMVQCKHMLFVSSLDQRLQVSEYELCKMTLLLSVPSKESLKHLEFLEIR----- 983
QY 377 LNAVDAIVYAKRSITSSQ--QDSMRILNL 406
DB 984 MYITKFL*KAIVKRR*NSQ*MD*RYQLTKIL 1015

```

```

APPLICANT: AHERNS-FATH, ISABELLE
: APPLICANT: BAENDLER, BERNARD
: TITLE OF INVENTION: HUMAN ANDROGEN RECEPTOR VARIANTS
: FILE REFERENCE: SCH-1794
: CURRENT APPLICATION NUMBER: US/09/997,267
: CURRENT FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: 60/255,078
: PRIOR FILING DATE: 2000-12-14
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 4
: LENGTH: 294
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-997-267-4

Query Match 19.2% Score 422.5; Db 9; Length 294;
Best Local Similarity 30.6%; Pred. No. 26-29;
Matches 96; Conservative 61; Mismatches 94; Indels 63; Gaps 9;

QY 61 EHTLPVNRKELKKKVSQNCASPVTPQSKRDHF-----CAVSDYASGYHGWSCG 115
DB 13 DHVLP-----DYEFHQKTCILGLDASG*HYGALITGS 47
QY 116 GVAFFKSTQSHNYTGATNCTIDKNPPKSCA*PLK*GV*W*W*G*TS*P*P*G*CV*PL 175
DB 18 GVAIFKAAQCFQ*GL*AS*SN*W*LL*GN*F*AN*DS*PL*EK*Y*EA*RT*LA*W*----- 98
QY 176 VPEQSADQELW*AK*AK*PS*CH*AP*VE*LE*-----L*MA*SP*GV*IT*TE*AF*P*H*V*IS* 240
DB 96 -KIK*Q*NI*KL*EP*P*AS*TS*P*ET*ET*KT*V*ST*P*Y*E*Q*P*IT*LA*E*IT*EG*V*AG 157
QY 231 -RISAPFTASMMMSITKLAK*ELVMS*AK*IP*GE*V*LS*P*V*RL*ES*W*P*V*IM*G 289
DB 158 HUNNGPSP*AL*LS*IN*IG*EK*Q*LV*W*KA*AL*P*E*P*NI*H*V*Q*MA*V*LY*SM*GL*W*FA 217
QY 290 LKMSLDHCK--L*H*AP*LV*LP-----K*E*G*V*E*LE*IT*H*ML*AI*SP*E*EL*KL*Q 340
DB 218 MGRPSTVNSPMLYFAPLVPEVPEKSPKPSQ*V*-----P*H*H*SP*E*Q*ML*Q*IT 269
QY 341 HKRYIVKAMILLN 354
DB 270 PGEFLCKMALLFS 283

RESULT 36
US-09-997-267-2
: Sequence 4, Application US/09997267
: Patent No. US20020165381A1
: GENERAL INFORMATION:
: APPLICANT: AHERNS-FATH, ISABELLE
: APPLICANT: BAENDLER, BERNARD
: TITLE OF INVENTION: HUMAN ANDROGEN RECEPTOR VARIANTS
: FILE REFERENCE: SCH-1794
: CURRENT APPLICATION NUMBER: US/09/997,267
: CURRENT FILING DATE: 2001-11-30
: PRIOR APPLICATION NUMBER: 60/255,078
: PRIOR FILING DATE: 2000-12-14
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 2
: LENGTH: 388
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-997-267-2

Query Match 19.2% Score 422.5; Db 9; Length 388;
Best Local Similarity 30.6%; Pred. No. 2,96-29;
Matches 96; Conservative 61; Mismatches 94; Indels 63; Gaps 9;

QY 61 EHTLPVNRKELKKKVSQNCASPVTPQSKRPDAH-----CAVSDYASGYHGWSCG 115
DB 13 DHVLP-----DYFPKPKTCILGLDASG*HYGALITGS 47

```


QY 402 LAMLMLSHR 413
DB 486 FAKLLRLPALR 497

RESULT 49

US-09-965-703-61
Sequence 61: Application US/09965703
Patent No. US2002019521A1
GENERAL INFORMATION:
APPLICANT: Rohm and Haas Company
APPLICANT: Patti, Subba Reddy
APPLICANT: Kapitskaya, Marianna Zinovjevna
APPLICANT: Gress, Dean Evelyn
TITLE OF INVENTION: No. US2002019521A1-1 Ecdysone Receptor Based Inhibitor Gene Expr
FILE REFERENCE: A010208
CURRENT APPLICATION NUMBER: US/09/965,703
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/191,355
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/256,799
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: PCT/US01/09050
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 61
LENGTH: 472
TYPE: PRT
ORGANISM: Choriostoneura fumiferana
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: No. US2002019521A1-1 Sequence
US-09-965-703-61

Query Match 18.78; Score 411.5; DB 10; Length 472;
Best Local Similarity 28.6%; Pred. No. 3,6628;
Matches 120; Conservative 68; Mismatches 136; Indels 95; Gaps 17.

QY 30 GCHSLPLVVRHQLSHIYARPKSP---WCFARSLPHLPVNPFTLRKKSQKAPASPVIG 86
DB 72 FCHSPSP-----SPPEKPDUTAMLDGLDDDTSPNPKNY-----PNH 110
QY 87 P-GSKRDHFCAGVCSYAGYGVWVSCGCKAFKRSIOGHNDYICPATNOCTIDKR 144
DB 111 PLSSSK--HICSGIGRAGSKHYGVYSGFCKGPFKPIVPMKI SYATCEENGTIDKQ 167
QY 145 KPSVQA-PIPKYTFVQWV-GSEFEKQGEIVPELSAFCQLH-CAVAKPSGCHAPVPE 204
DB 168 KNRQYGVYVCK-----LACNMPFPAVQDEPQENACCA-ETAHFSSVGVG----- 212
QY 205 LILHAI SPFOVLTLLEAPPHVILSPDSAPF-----TEASMSIILKAKR 251
DB 214 ---DELSTK--LIFMS-----LVAPSEHFOPLKVDGDSNVPKRYAPVSSIVQIDNK 262
QY 252 EIVHMLISWARKIPQFVHLSLHQVRLLESQWMEVLMGLKWRSDH----- 297
DB 263 QIALVYVWARDIPHFQI ELDQVYLKASWMLLEFALVMSMEYLFDERENGDTST 322
QY 298 FKKIIFARDVLDQFQVCEHIFETMLLAITS-SFRELKILHHHYLVKAMLLN 354
DB 323 TDQDLMQIMQMLHRSAD-QAGVCAIFQVLSLSEKPTLPMQVAYVAKAVILN 481
QY 355 SSMTPLVATQDSDSKRIAHINAVTDALVWVIAKSGISSQASMLAMLMLSHR 413
DB 482 PDVAGLKN-KQFVDVLRK--KMSCLDD-----YGRSRSRNEECRFASILLRLPALR 490

RESULT 40
US-09-922-226-3
Sequence 3: Application US/09922226
Publication No. US2003007766A1
GENERAL INFORMATION:

APPLICANT: Zhao, Yi
APPLICANT: Thacher, Scott M.
APPLICANT: Xiao, Jia-Hao
APPLICANT: Kusari, Jyotirmoy
APPLICANT: Chandraratna, Roshantha A.
TITLE OF INVENTION: Methods of Screening for Compounds that
Modulate Hormone Receptor Activity
FILE REFERENCE: P-AR 4681
CURRENT APPLICATION NUMBER: US/09/922,226
CURRENT FILING DATE: 2002-01-09
PRIOR APPLICATION NUMBER: US 60/284,797
PRIOR FILING DATE: 2001-04-18
NUMBER OF SEQ ID NOS: 191
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 463
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-226-3

Query Match 18.18; Score 403; DB 9; Length 463;
Best Local Similarity 31.9%; Pred. No. 26,27;
Matches 106; Conservative 54; Mismatches 141; Indels 32; Gaps 8;

QY 83 PVTGSKRDHFCAGVCSYAGYGVWVSCGCKAFKRSIOGHNDYICPATNOCTIDK 142
DB 127 PSTSPS-LVKHLCAGDRSSSKHYGVYSGCKGFFRTIKRLIYTCDKNGCTIDK 185
QY 143 NPKSGVCPPIPCYVW-VWYKSGREPCGYLVPQPSADQLP-VKAKPQCHAPV 202
DB 186 KQNRQYGVYVCKIWMQMPFPAVQDEPQENACCA-ETAHFSSVGVG-----TSCHDMP 235
QY 293 PELLTALFQVLTIL-LAEFTIVILSLISATFTASMMSTLAKELVHMSWAVY 262
DB 236 VERLEA---ELAVEKTESYGDMMENSTNDVY-----NCHAAKQDLFIVEMAKR 286
QY 263 ITCFVPLSTQVEIIFSCWMEVLMELMWSSTIHKGLIFATVILFEESKEVEGILE 322
DB 287 IPHSPDITLFDQVILLPACWMLLITASFHPSVSGVPIILATGLVHRSRASA-IVGS 345
QY 323 TFDMLLA-TTSRRELKIQHKEYLVKAMILLSSMYPLVLTQDSDSKRIAMLMNAV 381
DB 346 IFQVILIELVSKRMQKQKSELGLPAVLLENPAKQJNSPSFETLREKVATILEAVT 405
QY 482 DATVWVIAKSGISSQASMLAMLMLSHR 413
DB 406 KQ-----KYPEQSGFPAKLLRLPALR 427

Search completed: June 16, 2003, 13:41:52
Job time: 46.3127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 16, 2003, 12:52:26, Search time 23.043 seconds
(without alignments)
1731.646 Million cell updates/sec

Title: US-08-826-361A-6

Perfect score: 2203

Sequences: 1 MMSYSPSNVTNMRGHPROPT

QUSMRANIIMTSHVPAH 416

Sequence: B0208M62
Gapop 10.0, Gapext 0.5

Scored: 28224 seqs, 9614422 residues

Total number of hits satisfying chosen parameters: 28324

Minimum DB seq length: 0
Maximum DB seq length: 208000000

Post processing: Minimum Match 100%

Maximum Match 100%
Listing first 120 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2198	99.8	477	2 S71400	estrogen receptor
2	2198	99.8	530	2 J05939	estrogen receptor
3	1974	89.6	503	2 J06046	estrogen receptor
4	1075	48.8	589	1 ORCHE	estrogen receptor
5	1058	48.0	701	2 S64737	80K calcitonin recep
6	1056.5	48.0	620	2 T16423	estrogen receptor
7	1054	47.8	595	1 ORHUE	estrogen receptor
8	1050	47.7	599	1 ORHSE	estrogen receptor
9	1045	47.4	595	2 J47119	estradiol receptor
10	1044.5	47.4	586	1 ORX1E	estrogen receptor
11	1044	47.4	600	1 ORRTE	estrogen receptor
12	998	45.7	535	2 S68224	estrogen receptor
13	986.5	44.8	574	2 A37197	estrogen receptor
14	982.5	44.8	433	2 B03445	steroid hormone re
15	599	25.4	433	2 S58087	steroid hormone re
16	540.5	24.5	521	2 A29345	steroid hormone re
17	445	21.1	923	2 J53280	progesterone recep
18	444	20.6	930	2 A39506	progesterone recep
19	443	20.6	930	2 A39503	progesterone recep
20	449.5	20.4	758	2 S60586	glucocorticoid rec
21	448	20.3	933	1 ORHHP	glucocorticoid rec
22	440	20.0	776	1 S44047	glucocorticoid rec
23	437	19.8	919	2 A39248	androgen receptor
24	434	19.7	910	2 A34721	androgen receptor
25	434	19.7	911	2 B34721	androgen receptor
26	432.5	19.6	849	2 J02094	androgen receptor
27	431	19.6	771	2 A54273	glucocorticoid rec
28	430.5	19.5	984	2 A29512	glucocorticoid rec
29	430	19.5	786	2 A35466	progesterone recep

40	430	19.5	981	2 A41401	mineralocorticoid
41	428	19.4	902	2 B40944	androgen receptor
42	427	19.4	742	1 ORHHA	glucocorticoid rec
43	426.5	19.4	777	1 ORHHA	glucocorticoid rec
44	426	19.3	899	2 A55895	androgen receptor
45	418	19.0	783	1 A25691	glucocorticoid rec
46	417.5	19.0	533	2 S37781	retinoid X receptor
47	417.5	18.8	416	2 S-0663	retinoid acid recep
48	414.5	18.8	440	2 A34418	H-2 region II bind
49	414.5	18.8	448	2 D41727	retinoid X receptor
50	414.5	18.8	451	2 A41651	retinoid acid recep
51	414.5	18.8	520	2 B43718	RXR beta1 isoform
52	414	18.8	795	1 ORRIG	glucocorticoid rec
53	412	18.7	457	2 A43981	retinoid X receptor
54	411	18.7	462	2 S44490	RXR type hormone r
55	405	18.4	463	2 S26670	retinoid acid recep
56	405	18.4	463	2 B41727	retinoid X receptor
57	401.5	18.2	598	2 A37251	probable nuclear h
58	401	18.2	344	2 J51310	androgen receptor
59	399.5	18.1	563	2 J06023	nucleic growth facto
60	399	18.1	467	2 A47278	retinoid X receptor
61	395	17.9	603	1 ORMSN1	probable hormone r
62	393.5	17.9	462	1 S09592	retinoid X receptor
63	393	17.8	467	2 S26668	retinoid acid recep
64	391	17.7	470	2 B41977	retinoid receptor
65	388	17.6	385	2 J53158	orphan nuclear rec
66	388	17.6	441	2 J50515	retinoid X receptor
67	385	17.5	485	2 S48074	nuclear receptor 1
68	382.5	17.4	133	2 S26595	estrogen receptor
69	382	17.4	438	2 J150517	retinoid X receptor
70	381.5	17.3	577	2 S33763	hormone receptor N
71	380.5	17.3	495	2 A57053	stem cell nuclear
72	380	17.2	379	2 J50514	retinoid X recepto
73	376	17.1	452	2 S47633	RXR protein - A1r1
74	375	17.0	488	2 C41977	retinoid receptor
75	373.5	17.0	508	2 S11513	steroid hormone re
76	373.5	17.0	508	2 T13717	steroid hormone re
77	373	16.9	422	2 J50516	retinoid X recepto
78	365	16.6	628	2 J02493	nucleic derivedorp
79	364	16.5	598	2 A57040	T-cell nuclear rec
80	362	16.4	598	2 B46492	nuclear orphan rec
81	361	16.4	598	2 A46225	nuclear orphan rec
82	354	16.1	455	2 A36471	transcription fact
83	353.5	16.0	774	2 J06095	hepatocyte nuclear
84	353.5	16.0	625	2 S29330	nucleic derived rec
85	352	15.9	465	2 S22074	hepatocyte nuclear
86	351	15.9	465	2 S23502	hepatocyte nuclear
87	350.5	15.9	808	1 S33708	nucleic steroid/th
88	349	15.8	381	2 A46358	steroid/thyroid ho
89	347.5	15.8	504	2 J06095	hepatocyte nuclear
90	347	15.8	560	2 S27874	steroid hormone re
91	345.5	15.7	447	2 J18975	nuclear orphan rec
92	344	15.6	468	2 J04948	hepatocyte nuclear
93	344	15.6	468	2 J04947	hepatocyte nuclear
94	344	15.6	468	2 A56158	retinoid acid recep
95	343.5	15.5	448	2 B71133	apoptin protein A 1
96	342.5	15.5	414	2 J48975	apoptin protein A 1
97	342	15.5	459	2 A41377	retinoid acid recep
98	341.5	15.5	403	2 S35334	steroid hormone re
99	341.5	15.5	461	2 J04014	steroid hormone-mu
100	341.5	15.3	410	2 J50689	orphan receptor 50
101	341.5	15.3	453	2 J50674	retinoid acid recep
102	341.5	15.3	462	2 S05873	retinoid acid recep
103	340.5	15.3	464	2 A56558	retinoid acid recep
104	340.5	15.2	443	1 C35991	retinoid acid recep
105	339.5	15.2	466	2 S36218	hepatocyte nuclear
106	335	15.2	447	2 B34714	retinoid acid recep
107	333.5	15.1	445	2 A56043	steroid hormone re
108	333.5	15.1	454	1 A33903	retinoid acid recep
109	333.5	15.1	456	2 A34734	retinoid acid recep
110	333.5	15.1	462	1 A29491	retinoid acid recep
111	333	15.1	423	2 S02710	transcription fact
112	333	15.1	501	1 A56543	PI3-ET relatedorp

A:Accession: U09607
A:Molecule type: mRNA
A:Residues: 1-586 -Wt>
A:Cross references: chr12:9745
A:Comment: The steroid hormones and their receptors are involved in the regulation of complex cellular processes. Steroid hormone receptors are thought to be work as complex. In the absence of ligand, steroid hormone receptors are thought to be work as complex appears to recognize discrete DNA sequences upstream of transcription start sites. Steroid receptors belong to the superfamily of nuclear receptor proteins. They bind to specific DNA sequences and regulate gene expression by acting as transcription factors.
E:Keywords: DNA binding; nucleotides; phosphoprotein; steroid binding; steroid hormone receptor; transcription factor; zinc finger protein; zinc finger motif
F:1-174/Domain: amino-terminal /NH2
F:175-266/Domain: DNA binding #status Predicted_DNA
F:176-448/Domain: ERBA transforming protein homology - ERBA-
F:178-293/Region: zinc finger CCCC motif
F:214-287/Region: zinc finger CCCC motif
F:251-266/Region: nuclear location signal
F:292-544/Domain: steroid binding #status predicted_STH
F:180-181-187-205/Binding site: zinc (cys) #status predicted
F:216-223-232-235/Binding site: zinc (cys) #status predicted
F:231/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 47.4% Score 1044.5; DB 1; Length 586;
Best Local Similarity 50.7%; Pred. No. 1-20-70;
Mat. Ch. 5 210 230-241-250 71 Mismatch 10% Indels 27 Gaps 8;

22 SPNVLVPIPGHLSPLVNR-QLSHYAPQGSFWFAKSLFLILVNRIIKRVSINRR 80

[illegible]

Submitted to the EMBL data library, February 1990
 A:Reference number: S12464
 A:Accession: S12464
 A:Molecule type: mRNA
 A:Residues: 1-443, 776, 745-933 <KA2>
 A:Cross-references: EMBL:X51730; NID:935651; PDB:CA66018.1; PDB:945652
 R:Murphy, M.; Alder, M.; D'Aurilio, L.; Loosfelt, H.; Moriel, C.; Prilansky, E.; Guichet
 Blochem. Biophys. Res. Commun. 143, 740-748, 1987
 A:Title: Complete amino acid sequence of the human progesterone receptor deduced from cDNA
 A:Reference number: A03245; MUID:87184565; PMID:3551956
 A:Accession: A03245
 A:Molecule type: mRNA
 A:Residues: 1-225, 767, 227-355, 767, 257-659, 767, 661-933 <MS>
 A:Cross-references: GB:M15716; NID:9189934; PDB:AAA60081.1; PDB:9189935
 C:Keywords: C:Keywords: progesterone receptor; cDNA; transforming protein homology
 A:Map position: 11922, 1-11922, 3
 A:Superfamily: progesterone receptor; DNA binding; nucleus; phosphoprotein; steroid hormone
 C:Keywords: alternative splicing; DNA binding; nucleus; phosphoprotein; steroid hormone
 E1-933/Protein: progesterone receptor form B #status predicted <MA1>
 E1-165/933/Protein: progesterone receptor form A #status predicted <MA2>
 E1-565/829/Domain: cDNA transformed protein homology <EPA>
 E1-567/587/Region: zinc finger core motif
 E1-603-627/Region: zinc finger core motif
 E1-601-933/Domain: steroid binding #status predicted <STB>
 E1-1740/Protein: steroid binding site (phosphate) #status predicted
 E1-277-292/529/Protein: steroid binding site (phosphate) #status predicted
 E1-329-374/601/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 20.38, Score 448; DB 1; Length 933;
 Best local similarity 32.08, Prod. No. 7, 20-29;
 Matches 107; Conservative 63; Mismatches 127; Indels 28; Gaps 10;

QY 56 GAVSTVAVYVWVSTPTFAFFPSTQHNQVAVANQVTHV NPEKSGVQVPE 155
 DB 567 CLINDFASGCHVAVITGSGCKVEFKPAMEGQNHVYICAGPNDCLVQK IRRKNVAPVQVPE 626
 QY 156 GYEVWVWVQSPRPVQVPIVPPQSPAPDQ--HCAKAPKSGSHAPVVEILLALSP 213
 DB 627 CTGAGMVLGGRKFKK--ENKRVVVALDAVALPQVLPVNSQALSGP-----TESPQ 678
 QY 214 Q-----LVTLLEAPPHVITSPSA-PEPEASMMSTKLAKELVMIISMAKIKQ 245
 DB 679 QHIGITPPPLNLMISIEDVYAGHCKTKPOTSSLLITSLNOLGKROLISVVMKNSLPG 748
 QY 246 FVPI SLPGVPIESQWMEVIMGI MNPSTIDHPG--ALFAPVVLVDPVQKVEGILEI 723
 DB 749 FENHIDPQITLLGYSMVSLNPGIOWPSYKHSQGLVYAPVPLINEQPMK--ESSFYSL 797
 QY 424 FDMILATTSREKELKQKEYLQVAKMLINS-SMPLVATQV-DANSSKRLHLINAV 481
 DB 748 CTIMQITQVPEVKIQVSPFTQKVVITLNTIHLPELPSQVPEHPSVYIPELFA-- 855
 QY 482 DALVWVAKSGISSQSSMLANLL 406
 DB 856 --LGLRQKGVSSSSQRFYQTLKLL 877

RESULT 22
 S41 47
 glucocorticoid receptor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 14-Jan-1995 #sequence revision 27 07-11-1995 #local database 22 10 1999
 C:Accession: S45348; S44047; S42083
 R:Adoo, X.; Kalkhoven, E.; Petersen-Maduro, J.; van der Burg, B.; Westrey, C.H.J.;
 Blochem. Biophys. Acta 1218, 194-198, 1994
 A:Title: Expression of the glucocorticoid receptor gene is regulated during early embry-
 A:Reference number: S45348; MUID:94289478; PMID:8018720
 A:Accession: S45348
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA

A:Residues: 1-776 <GAW>
 A:Cross-references: EMBL:X72211
 A:Experimental source: embryonic stage 17
 R:Gao, X.; Kalkhoven, E.; Petersen-Maduro, J.; van der Burg, B.; Westrey, C.H.J.;
 Submitted to the EMBL data library, May 1993
 A:Description: The expression of glucocorticoid receptor gene is regulated during ear-
 A:Reference number: S44047
 A:Accession: S44047
 A:Molecule type: mRNA
 A:Residues: 1-271, 767, 273-776 <GAW>
 A:Cross-references: EMBL:X72211; NID:944042; PDB:CA651010.1; PDB:944044
 A:Experimental source: embryonic stage 17
 R:Gao, X.
 Submitted to the EMBL data library, February 1994
 A:Reference number: S42083
 A:Accession: S42083
 A:Molecule type: mRNA
 A:Residues: 152-521, 767, 503, 767, 505, 767, 507-510, 767, 512-518, 767, 520-543, 767, 545-550, 767
 A:Cross-references: EMBL:X77764; NID:9456373; PDB:CA654804.1; PDB:9456374
 A:Experimental source: liver
 C:Superfamily: glucocorticoid receptor; cDNA; transforming protein homology
 C:Keywords: DNA binding; nucleus; steroid hormone receptor; transcription regulation
 E1-418-473/Domain: cDNA transformed protein homology <EPA>
 E1-120-143/Region: zinc finger core motif
 E1-170-180/Region: zinc finger core motif

Query Match 20.08, Score 440; DB 1; Length 776;
 Best local similarity 32.08, Prod. No. 1, 46-27;
 Matches 120; Conservative 55; Mismatches 146; Indels 44; Gaps 14;

QY 65 FVNPETLKKVSGNPGAS-----VTPGSKRIAHFVAVSGVAVSGVHG 109
 DB 676 FVNPVTRFVVI SGVSPGIRLSVSPSSLSISPPPR-----LGLVSGPASCQVTE 433
 QY 110 VWSDEGKAPFKESLQGHNDYLPAINQVLIHANEKSSVAVALEPVEVWVWVQSRPE 169
 DB 434 VITGSGTQVPEKPVVGHVNLZAGPNCILQKPNVNCVQVPEVPLDGMNIEARRP 492
 QY 170 RQGVLYVRQSAIEQLHCNCAKAPKSGSHAPVVEILLALSPQVLTLLFAPIR 226
 DB 493 FKKIKIQDQSTVA-----TAPSEPTSMPTIVPASNVAQITP--TITSTIVTPPE 541
 QY 227 VVIS--PEPAPPTASMMSTKLAKELVMIISMAKIPVEELSLFQVPLIESVME 284
 DB 542 VVISGYSSTIDPTTRPLMSLSNMLGSPGVVAVVMAKATGPNHIDMQMTLLQVSMW 601
 QY 285 VIMMGLMNPSTIDHPK--LIFAPVLDVDP-----EGKGVSEILEIPMLATTSPK 335
 DB 602 LMVFALMWSYKGTGNSLTYFAPDVTVEDEGMHLPFQEPQGE-----MKIAGENS 653
 QY 336 FIKIGIKRYIVKAMITLNSMVPVYATQVAVSPETAH--TNAVTVAVVWVAKNSI 393
 DB 654 SLQTSYDQVLEKVKLL-----MCTTPKRLKSHALPELPAIVIKELAKVAKRKN 706
 QY 394 SSQ--QVSMRLANLL 406
 DB 707 SSQNMQRVQYQTLKLL 721

RESULT 24
 A39248
 androgen receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 01 Oct 1993 #sequence revision 34 Oct 1993 #local database 24 Nov 1999
 C:Accession: A39248; A39328; A41673; A60346; A19412; A37553; A41699; A44944; A42244;
 R:Kubundin, D.H.; Acad, S.G.; Brown, T.R.; Stancovski, I.A.; Huggs, B.N.; Mison, C.J.; Wilson, E.M.;
 Proc. Natl. Acad. Sci. U.S.A. 90, 9534-9538, 1993
 A:Title: Sequence of the intron/exon junctions of the coding region of the human and-
 A:Reference number: A39248; MUID:95083302; PMID:2594784
 A:Accession: A39248
 A:Molecule type: DNA
 A:Residues: 1-919 <GB>
 A:Cross-references: GB:M27423; GB:M27430; NID:9178904; PDB:AAA1886.1; PDB:9178906

GenScore version 5.1.6
Copyright (c) 1993 - 2003 CompuLink Ltd.

GM protein - protein search, using SW model

Run on: June 16, 2003, 07:49:07 ; Search time 10.5242 seconds
(without alignments)
1639,467 Million cell updates/sec

Title: US-08-826-361A-6

Sequence: 1 MMYSLPSNVNMLEGAPSPOT.....QNSPLANLMLSHVHAP 414

Search: BLAST/BLAST2
Gapop 10.0 ; Gapext 0.5

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 120 summaries

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2198	99.8	540	ESR2_HUMAN	Q62773 home sapien
2	2064	99.2	540	ESR2_FALTA	Q62773 home sapien
3	1981	89.4	540	ESR2_PAT	Q62773 home sapien
4	1980.5	89.9	527	ESR2_SHEEP	Q62773 home sapien
5	1978	89.8	530	ESR2_MOUSE	Q62773 home sapien
6	1973.5	89.6	527	ESR2_BOVIN	Q62773 home sapien
7	1928	87.5	526	ESR2_PIG	Q62773 home sapien
8	1762	80.0	472	ESR2_PORC	Q62773 home sapien
9	1756.5	79.7	554	ESR2_STROM	Q62773 home sapien
10	1750	79.4	472	ESR2_CHICK	Q62773 home sapien
11	1413.5	59.6	573	ESR2_ANGIA	Q62773 home sapien
12	1243	56.4	568	ESR1_CANAV	Q62773 home sapien
13	1247	56.2	568	ESR2_WYAT	Q62773 home sapien
14	1243	56.0	673	ESR2_MITON	Q62773 home sapien
15	1236.5	55.7	575	ESR2_LCTTU	Q62773 home sapien
16	1218	55.3	610	ESR2_CANAV	Q62773 home sapien
17	1194.5	54.2	559	ESR2_SPAUV	Q62773 home sapien
18	1174	53.5	545	ESR2_MITON	Q62773 home sapien
19	1164	52.8	557	ESR2_ORENI	Q62773 home sapien
20	1150	52.2	279	ESR2_MACRO	Q62773 home sapien
21	1070	48.6	589	ESR1_CHICK	Q62773 home sapien
22	1069	48.5	589	ESR1_PORC	Q62773 home sapien
23	1064	48.3	595	ESR1_HUMAN	Q62773 home sapien
24	1062	48.2	585	ESR1_ORENI	Q62773 home sapien
25	1056.5	48.0	620	ESR1_SPAUV	Q62773 home sapien
26	1056.5	48.0	620	ESR1_ORENI	Q62773 home sapien
27	1056.5	48.0	620	ESR1_ORENI	Q62773 home sapien
28	1048	47.6	569	ESR1_PAPAE	Q62773 home sapien
29	1044	47.4	581	ESR1_PAPAE	Q62773 home sapien
30	1045	47.4	595	ESR1_PAPAE	Q62773 home sapien
31	1044.5	47.4	586	ESR1_XENLA	Q62773 home sapien
32	1044.5	47.4	595	ESR1_XENLA	Q62773 home sapien
33	1044	47.4	600	ESR1_PAT	Q62773 home sapien
34	1043	47.3	525	ESR1_MITON	Q62773 home sapien
35	1042.5	47.3	594	ESR1_HORSE	Q62773 home sapien
36	1035	47.0	583	ESR1_ORENI	Q62773 home sapien
37	1019	46.3	617	ESR1_LCTTU	Q62773 home sapien
38	1007.5	45.7	622	ESR1_GNCPY	Q62773 home sapien
39	998	45.3	535	ESR1_SALSA	Q62773 home sapien
40	852	38.7	307	ESR1_CNEUN	Q62773 home sapien
41	654	29.7	349	ESR1_ANGIA	Q62773 home sapien
42	573.5	26.0	458	ESR3_HUMAN	Q62773 home sapien
43	572.5	26.0	433	ESR2_PAT	Q62773 home sapien
44	572	26.0	433	ESR2_HUMAN	Q62773 home sapien
45	559	25.4	433	ESR2_MOUSE	Q62773 home sapien
46	539.5	24.5	519	ESR1_HUMAN	Q62773 home sapien
47	535	24.1	462	ESR1_MOUSE	Q62773 home sapien
48	465	21.1	923	ESR1_PAT	Q62773 home sapien
49	464	21.1	923	ESR1_MOUSE	Q62773 home sapien
50	456	20.7	377	ESR1_SHEEP	Q62773 home sapien
51	453	20.6	930	ESR1_PAT	Q62773 home sapien
52	443.5	20.4	758	ESR1_PAT	Q62773 home sapien
53	448	20.3	933	ESR1_HUMAN	Q62773 home sapien
54	440	20.0	776	ESR1_XENLA	Q62773 home sapien
55	438	19.9	895	ESR1_PAPAE	Q62773 home sapien
56	438	19.9	895	ESR1_PAPAE	Q62773 home sapien
57	438	19.9	911	ESR1_PAPAE	Q62773 home sapien
58	437	19.8	919	ESR1_PAPAE	Q62773 home sapien
59	435.5	19.8	703	ESR1_PAT	Q62773 home sapien
60	435	19.7	709	ESR1_PAT	Q62773 home sapien
61	433	19.7	776	ESR1_PAT	Q62773 home sapien
62	433	19.7	884	ESR1_PAT	Q62773 home sapien
63	430.5	19.5	174	ESR1_PAT	Q62773 home sapien
64	430.5	19.5	984	ESR1_PAT	Q62773 home sapien
65	430	19.5	786	ESR1_PAT	Q62773 home sapien
66	430	19.5	981	ESR1_PAT	Q62773 home sapien
67	428	19.4	922	ESR1_PAT	Q62773 home sapien
68	427.5	19.4	777	ESR1_PAT	Q62773 home sapien
69	427.5	19.4	777	ESR1_PAT	Q62773 home sapien
70	427.5	19.4	777	ESR1_PAT	Q62773 home sapien
71	426.5	19.4	777	ESR1_PAT	Q62773 home sapien
72	426	19.3	899	ESR1_PAT	Q62773 home sapien
73	425.5	19.3	807	ESR1_PAT	Q62773 home sapien
74	425	19.3	74	ESR1_PAT	Q62773 home sapien
75	424	19.2	798	ESR1_PAT	Q62773 home sapien
76	422	19.2	721	ESR1_PAT	Q62773 home sapien
77	418	19.0	783	ESR1_PAT	Q62773 home sapien
78	417.5	19.0	533	ESR1_PAT	Q62773 home sapien
79	417.5	19.0	458	ESR1_PAT	Q62773 home sapien
80	414.5	18.8	529	ESR1_PAT	Q62773 home sapien
81	414	18.8	795	ESR1_PAT	Q62773 home sapien
82	413	18.7	461	ESR1_PAT	Q62773 home sapien
83	412	18.7	467	ESR1_PAT	Q62773 home sapien
84	411.5	18.7	472	ESR1_PAT	Q62773 home sapien
85	411	18.7	462	ESR1_PAT	Q62773 home sapien
86	411	18.7	612	ESR1_PAT	Q62773 home sapien
87	409	18.6	977	ESR1_PAT	Q62773 home sapien
88	405	18.4	463	ESR1_PAT	Q62773 home sapien
89	403	18.3	463	ESR1_PAT	Q62773 home sapien
90	401.5	18.2	598	ESR1_PAT	Q62773 home sapien
91	399.5	18.1	597	ESR1_PAT	Q62773 home sapien
92	398	18.1	467	ESR1_PAT	Q62773 home sapien
93	395	17.9	601	ESR1_PAT	Q62773 home sapien
94	393.5	17.9	462	ESR1_PAT	Q62773 home sapien
95	394	17.8	467	ESR1_PAT	Q62773 home sapien
96	392.5	17.8	598	ESR1_PAT	Q62773 home sapien
97	392	17.7	473	ESR1_PAT	Q62773 home sapien
98	388	17.6	485	ESR1_PAT	Q62773 home sapien
99	388	17.6	441	ESR1_PAT	Q62773 home sapien
100	385	17.6	495	ESR1_PAT	Q62773 home sapien
101	385	17.5	485	ESR1_PAT	Q62773 home sapien
102	381.5	17.3	577	ESR1_PAT	Q62773 home sapien
103	381	17.3	488	ESR1_PAT	Q62773 home sapien
104	380.5	17.3	495	ESR1_PAT	Q62773 home sapien
105	375.5	17.0	508	ESR1_PAT	Q62773 home sapien
106	373.5	17.0	508	ESR1_PAT	Q62773 home sapien

107	470	16.8	111	1	ESR1_SHEEP	P49885 ovls atces
108	465	16.6	628	1	NK43_RAT	p51179 ratfus novy
109	464	16.5	598	1	NK42_HUMAN	p44454 hemo sapien
110	463.5	16.5	480	1	NK61_HUMAN	Q15406 hemo sapien
111	464	16.5	626	1	NK43_HUMAN	Q12570 hemo sapien
112	462	16.4	496	1	NK21_MOUSE	Q19418 ratfus novy
113	462	16.4	598	1	NK42_RAT	Q07917 ratfus novy
114	461	16.4	598	1	NK42_MOUSE	Q05219 mus musculu
115	459.5	16.3	408	1	ETFB_MOUSE	Q05192 atropophila
116	456.5	16.2	408	1	HN4G_MOUSE	Q19606 mus musculu
117	451	15.9	465	1	HN4A_MOUSE	p22449 ratfus novy
118	451	15.9	465	1	HN4A_RAT	Q14541 hemo sapien
119	449.5	15.9	408	1	HN4G_HUMAN	p45448 mus musculu
120	447	15.8	560	1	NK52_MOUSE	

ALIGNMENTS

RESULT 1

ESR2_HUMAN STANDARD: PRT: 510 AA

AC G92711: 0606087; 060608; 075584; 060702; 060703; 075583; G910K9;

AC G908V6: G908D3;

AT 15 JUL 1999 (Ref: 48, Created)

AT 15 JUL 1999 (Ref: 48, Last sequence update)

AT 16 OCT 2001 (Ref: 40, Last annotation update)

DE Estrogen receptor beta (ER-beta).

GN ESR2 OR NR4A2 OR ESTRB.

OS Homo sapiens (human).

OR Biocytol: Metazoa; Chordata; Vertebrata; Euteleostomi;

OR Mammalia; Eularchia; Primates; Catartini; Hominidae; Homo.

OX NCBI TaxID 9606;

PN 111

RP SEQUENCE FROM N.A. (ISOFORM BETA-1).

RX MEDLINE 98119878; PubMed 947491;

RA Obara S., Inoue S., Matsuda T., Hiroi H., Ohtano A., Hosoi T., Ouchi Y., Muramatsu M.;

RA "The complete primary structure of human estrogen receptor beta (ERbeta) and its heterodimerization with ER alpha in vivo and in vitro.";

RT Biochem. Biophys. Res. Commun. 241:122-126(1998).

RL

RN 121

RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.

RX TISSUE Testis, and breast;

EX MEDLINE 98002867; PubMed 9606672;

RA Moore L.T., McKee D.D., Stewart-Koster K., Moore L.B., Jones S.A., Horne L.L., Su J., Lee K.L., Stewart S.A., Johnson J.M., Willison T.M.;

RA "Cloning and characterization of human estrogen receptor beta isoforms.";

RT Biochem. Biophys. Res. Commun. 247:75-79(1998).

RL

RN 131

RP SEQUENCE OF AN 510 FROM N.A. (ISOFORM BETA-1), AND CHARACTERIZATION.

RX TISSUE Testis;

EX MEDLINE 96454875; PubMed 8749433;

RA Mosselman S., Bolman J., Dijkema R.;

RA "ER beta: identification and characterization of a novel human estrogen receptor.";

RT FEBS Lett. 392:49-53(1996).

RL

RN 141

RP SEQUENCE FROM N.A. (ISOFORM BETA 2/5X), AND CHARACTERIZATION.

RX TISSUE Testis;

EX MEDLINE 9847008; PubMed 9671411;

RA Obara S., Inoue S., Matsuda T., Ouchi Y., Hosoi T., Ouchi Y., Muramatsu M.;

RA "Molecular cloning and characterization of human estrogen receptor beta cDNA: a potential inhibitor of estrogen action in human.";

RT Nucleic Acids Res. 26:4505-4512(1998).

RL

RN

RP SEQUENCE OF 59 510 FROM N.A. (ISOFORMS BETA 5A AND BETA-6).

RX TISSUE Endometrium;

RA Branderhorst A.W., Lebevie D., Taylor R.N., Datto R.B.;

RA Submitted (JUN 1998) to the EMBL/Genbank/DBJ databases.

KN [61

RP SEQUENCE FROM N.A. (ISOFORM BETA-2A).

RT TISSUE ovary;

EX MEDLINE 98448489; PubMed 9685228;

RA La B., Loyee E., DeGraw H., Murphy L.C., Watson E.H.;

RA "Estrogen receptor-beta mRNA variants in human and murine tissues.";

RL Mol. Cell. Endocrinol. 138:199-204(1998).

RN 171

RP SEQUENCE OF 1 69 FROM N.A.

RA La B.C., Ludwig R.;

RA "Cloning and characterization of the estrogen receptor beta gene promoter.";

RT Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.

RL

RN 181

RP CHARACTERIZATION.

RX MEDLINE 97467483; PubMed 9425415;

RA Favo F., Taylor J., Santharalingam S., Coombes R.C., Ali S.;

RA "Human estrogen receptor beta binds DNA in a manner similar to and dimerizes with estrogen receptor alpha.";

RL J. Biol. Chem. 272:26832-26838(1997).

CC 1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF ER ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN DEPENDENT MANNER.

CC ISOFORM BETA-EX LACKS LIGAND BINDING ABILITY AND HAS NO OR ONLY VERY LOW ERE BINDING ACTIVITY RESULTING IN THE LOSS OF LIGAND-DEPENDENT TRANSACTIVATION ABILITY. DNA BINDING BY ER-ALPHA AND BETA IS RAPIDLY LOST AT 47 DEGREES CELSIUS IN THE ABSENCE OF LIGAND WHILE IN THE PRESENCE OF 17 BETA-ESTRADIOL AND 4 HYDROXY TAMOXIFEN LOSS IN DNA BINDING AT ELEVATED TEMPERATURE IS MORE GRADUAL.

CC 1- SUBUNIT: BINDS DNA AS A HOMODIMER, CAN FORM A HETERODIMER WITH ER ALPHA. ISOFORM BETA-2/5X PREFERENTIALLY FORMS A HETERODIMER WITH ER-ALPHA RATHER THAN ER-BETA AND INHIBITS DNA BINDING BY ER ALPHA.

CC 1- SUBCELLULAR LOCATION: Nucleus.

CC 1- ALTERNATIVE PRODUCTS: 8 ISOFORMS: BETA-1 (SHOWN HERE), BETA 2/5X, BETA-2A, BETA-3, BETA-4, BETA-5, BETA 5A AND BETA-6; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC 1- TISSUE SPECIFICITY: ISOFORM BETA-1 IS EXPRESSED IN TESTIS AND OVARY, AND AT A LOWER LEVEL IN HEART, BRAIN, PLACENTA, LIVER, SKELETAL MUSCLE, SPLEEN, THYMUS, PROSTATE, COLON, BONE MARROW, MAMMARY GLAND AND UTERUS. ALSO FOUND IN UTERINE BONE, BREAST, AND OVARIAN TUMOR CELL LINES, BUT NOT IN COLON AND LIVER TUMORS.

CC ISOFORM BETA-2 IS EXPRESSED IN SPLEEN, THYMUS, TESTIS AND OVARY AND AT A LOWER LEVEL IN SKELETAL MUSCLE, PROSTATE, COLON, SPALL, INTERSTINE, LEUKOCYTES, BONE MARROW, MAMMARY GLAND AND UTERUS.

CC ISOFORM BETA-3 IS FOUND IN TESTIS. ISOFORM BETA-4 IS EXPRESSED IN TESTIS, AND AT A LOWER LEVEL IN SPLEEN, THYMUS, OVARY, MAMMARY GLAND AND UTERUS. ISOFORM BETA-5 IS EXPRESSED IN TESTIS, PLACENTA, SKELETAL MUSCLE, SPLEEN AND LEUKOCYTES, AND AT A LOWER LEVEL IN HEART, LUNG, LIVER, KIDNEY, PANCREAS, THYMUS, PROSTATE, COLON, SPALL, INTERSTINE, BONE MARROW, MAMMARY GLAND AND UTERUS. NOT EXPRESSED IN BRAIN.

CC 1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.

CC 1- STIMULATORY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC NRS SUBFAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed, misuse by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/submit/submit.html> or send an email to license@isb.scrib.ch).

DR EMBL: AB006590; JAA24953.1;

DR EMBL: AF051427; AAC05985.1;

DR EMBL: AF051428; AAC05251.1;

DR EMBL: AF061054; AAC49784.1;

DR EMBL: AF061055; AAC49785.1;

DR EMBL: AF060555; AAC15243.1;

DR EMBL: X99101; CAA67555.1; ALT_INT.

DR EMBL: AB006589; JAA11966.1;

```

DR EMBL: AF074548; AAC25602.1; -
DR EMBL: AF074549; AAC25603.1; -
DR EMBL: AF124790; AAC12580.1; -
DR EMBL: AF047446; AAC1786.1; -
DR EMBL: AF191544; AAC24242.1; -
DR TRANSFAC: T04651; -
DR Gene: HMG; 4468; ESR2.
DR MIM: 601663; -
DR InterPro: IPR000536; Hormone_rec_19.
DR InterPro: IPR001724; Steroid_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zt-C4_1.
DR PRINTS: PR00198; STDRHORMONER.
DR PRINTS: PR00047; STROIDPTNGR.
DR Prodom: PD000035; Znf_C4steroid; 1.
DR SMART: SM00399; Znf_C4_1.
DR SMART: SM00399; Znf_C4_1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation
FT DOMAIN 1 148 MODULATING.
FT DNABIND 149 214 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 149 169 C4-TYPE.
FT ZN_FING 185 209 C4-TYPE.
FT DOMAIN 215 530 STEROID-BINDING.
FT MOD_RES 87 87 PHOSPHORYLATION (BY SIMILARITY).
FT VASPLIC 488 488 PHOSPHORYLATION (BY SIMILARITY).
FT VASPLIC 419 423 IVER-1; M2NA (IN ISOFORM BETA-2A).
FT VASPLIC 319 409 MISSING (IN ISOFORM BETA-5A).
FT VASPLIC 424 530 MISSING (IN ISOFORM BETA-2A).
FT VASPLIC 365 375 DECKVEGLE -> YVPSHSDPGC (IN ISOFORM BETA-6).
FT VASPLIC 476 530 MISSING (IN ISOFORM BETA-6).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> VASASQTLSECYMM
FT VASPLIC 469 530 ETLPEATMEO (IN ISOFORM BETA-2/5X).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> SLSLSMELMLREAS
FT VASPLIC 469 530 CHRVPTPTGAAHMSVSPSPFACQUPPE (IN ISOFORM BETA-3).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> EWCCKQELHKLIS
FT VASPLIC 469 530 (IN ISOFORM BETA-4).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> EYAF (IN ISOFORM BETA-5).
FT VASPLIC 469 530 F -> G (IN REF. 3).
FT SPANPNE 530 AA; 54216 MW; 86AE3421599241BA CPO64;
Query Match 99.8%; Score 2198; DB 1; Length 540;
Host Local Similarity 100.0%; Prod. No. 5.7e-181;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DR EMBL: AF074548; AAC25602.1; -
DR EMBL: AF074549; AAC25603.1; -
DR EMBL: AF124790; AAC12580.1; -
DR EMBL: AF047446; AAC1786.1; -
DR EMBL: AF191544; AAC24242.1; -
DR TRANSFAC: T04651; -
DR Gene: HMG; 4468; ESR2.
DR MIM: 601663; -
DR InterPro: IPR000536; Hormone_rec_19.
DR InterPro: IPR001724; Steroid_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zt-C4_1.
DR PRINTS: PR00198; STDRHORMONER.
DR PRINTS: PR00047; STROIDPTNGR.
DR Prodom: PD000035; Znf_C4steroid; 1.
DR SMART: SM00399; Znf_C4_1.
DR SMART: SM00399; Znf_C4_1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation
KW Zinc-finger; Steroid-binding; Alternative splicing; Phosphorylation
FT DOMAIN 1 148 MODULATING.
FT DNABIND 149 214 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 149 169 C4-TYPE.
FT ZN_FING 185 209 C4-TYPE.
FT DOMAIN 215 530 STEROID-BINDING.
FT MOD_RES 87 87 PHOSPHORYLATION (BY SIMILARITY).
FT VASPLIC 488 488 PHOSPHORYLATION (BY SIMILARITY).
FT VASPLIC 419 423 IVER-1; M2NA (IN ISOFORM BETA-2A).
FT VASPLIC 319 409 MISSING (IN ISOFORM BETA-5A).
FT VASPLIC 424 530 MISSING (IN ISOFORM BETA-2A).
FT VASPLIC 365 375 DECKVEGLE -> YVPSHSDPGC (IN ISOFORM BETA-6).
FT VASPLIC 476 530 MISSING (IN ISOFORM BETA-6).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> VASASQTLSECYMM
FT VASPLIC 469 530 ETLPEATMEO (IN ISOFORM BETA-2/5X).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> SLSLSMELMLREAS
FT VASPLIC 469 530 CHRVPTPTGAAHMSVSPSPFACQUPPE (IN ISOFORM BETA-3).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> EWCCKQELHKLIS
FT VASPLIC 469 530 (IN ISOFORM BETA-4).
FT VASPLIC 469 530 SKEWERP LMKCKNVVPYDILHMLNAHVLGKSTLIG
FT VASPLIC 469 530 SECSPADSKSKESONPOSO -> EYAF (IN ISOFORM BETA-5).
FT VASPLIC 469 530 F -> G (IN REF. 3).
FT SPANPNE 530 AA; 54216 MW; 86AE3421599241BA CPO64;
Query Match 99.8%; Score 2198; DB 1; Length 540;
Host Local Similarity 100.0%; Prod. No. 5.7e-181;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


PT	PRO. RES.	487	488	488																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													</
----	-----------	-----	-----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

[illegible]

DR HSPD: P03372; HFO.
 DR InterPro: IPR000536; Hormone_rec_Lig.
 DR InterPro: IPR01628; Zn_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; z1-c4_1.
 DR PRINTS: PR00047; STEROIDFINDER.
 DR PRODOM: P000035; Zn_C4steroid_1.
 DR SMART: SM00430; ZNF_C4_1.
 DR SMART: SM00399; ZNF_C4_1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding.
 FT DOMAIN 1 104 MODULATING.
 FT DNA_BIND 105 170 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 105 125 CA TYPE.
 FT ZN_FING 141 165 CA TYPE.
 FT DOMAIN 171 472 STEROID-BINDING.
 ST SEQUENCE 472 AA; 53411 MW; DF7A78F0E1BD1BHD CMC64;
 Query Match 80.0%; Score 1762; ID 1; Length 472;
 Best Local Similarity 79.8%; Pred. No. 1,3e-143;
 Matches 432; Conservative 56; Mismatches 42; Indels 5; Gaps 5;
 QY 1 MNYSTPSNVTNLEGGPGKQTTSPNVLWPTPCHSLVVERQLSHLYAFKSTWQKASL 60
 DB 9 MNVNIASNGDSASVROISPSISLWNAHSHSLHQLSILYAFKSTWQKAPPL 66
 QY 61 EHTLVNPKETLKPKKSNQWCAAPV-TGPSKSTAHFVAV-SYVNSVYVWVSPVYKAF 119
 DB 69 EHTLVNPKETLKPKKSNQWCAAPV-TGPSKSTAHFVAV-SYVNSVYVWVSPVYKAF 129
 QY 120 FRSSTLGHVYLTPLNCTTTPKPFSTLALTEPTEVAVNGVNGFEETSELEHPE 179
 DB 129 FRSSTLGHVYLTPLNCTTTPKPFSTLALTEPTEVAVNGVNGFEETSELEHPE 194
 QY 189 KNSFD--CMGKTKYVNAFTVKEILLSTVSPQGVETLLFAEPNVI VSPSPKPTFA 245
 DB 240 SMNMSITKLADKELVHMTSMARKTGFVFLSTFGVRLLESTVMGVLMGIMWSTIDPG 259
 QY 245 SMNMSITKLADKELVHMTSMARKTGFVFLSTFGVRLLESTVMGVLMGIMWSTIDPG 305
 DB 300 KLIFADVLDRDFGKCVGEILEFDMLATTSRELEKLOKREYLVKAMLLNSSKVP 359
 QY 305 KLIFADVLDRDFGKCVGEILEFDMLATTSRELEKLOKREYLVKAMLLNSSKVP 445
 DB 446 LVTATLVANSSPRLAHINAVTDAVWVAWSQTSNQQSMPLMILSLHVGVA 415
 DB 466 LVTATLVANSSPRLAHINAVTDAVWVAWSQTSNQQSMPLMILSLHVGVA 419
 RESULT 9
 ESR2_STDVH STANDARD: PRI: 554 AA.
 AC GAFVYV.
 DT 16-001-2001 (Rel. 40, created)
 DT 16-001-2001 (Rel. 40, last sequence update)
 DT 16-001-2001 (Rel. 40, last annotation update)
 DE Estrogen receptor beta (ER-beta)
 GN ESR2 OR NK42.
 OS Steroid receptors (Steroids)
 OC Eukaryota; Metazoa; Chordata; Gracilaria Vertebrata; Euteleostomi;
 OC Actinopterygii; Aves; Neognathae; Passeriformes; Strigidae; Strigus
 OC NCB1; taxid 9172;
 RN [1]
 RP SOURCE: FROM N.A.
 RC TISSUE: Brain.
 FX MEDLINE: 99427866; PubMed: 10499520;
 RA Reutimard D.J., Bentley G.E., Ballhazart J., Turek F.W., Ball G.P.;
 RT "Androgen receptor, estrogen receptor alpha, and estrogen receptor
 RI beta show distinct patterns of expression in rodent and human
 RI nuclei of European starlings.";

RL Inductinology 140:4633-4641(1999).
 CC -1- FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF ER
 CC ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 CC ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN-DEPENDENT MANNER.
 CC EREALLY SYNTHESIZED ESTROGENS MAY ACT VIA ER BETA, IN ADDITION TO
 CC ER ALPHA, TO MEDATE SEASONAL OR DEVELOPMENTAL EFFECTS ON NEARBY
 CC SONG MOTIL.
 CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER, CAN FORM A HETERODIMER WITH ER-
 CC ALPHA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- TISSUE SPECIFICITY: BRAIN, PITUITARY, SKELETAL MUSCLE, LIVER,
 CC ADRENAL, KIDNEY, INTESTINE, AND OVARY.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation;
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre/infocentre.do>
 CC or send an email to license@ebi.ac.uk).
 DR HMR1: AF11513; AAD56594.1;
 DR HSPD: P03372; HFO.
 DR InterPro: IPR000536; Hormone_rec_Lig.
 DR InterPro: IPR01628; Steroid_receptor.
 DR InterPro: IPR01628; Zn_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; z1-c4_1.
 DR PRINTS: PR00047; STEROIDFINDER.
 DR PRODOM: P000035; Zn_C4steroid_1.
 DR SMART: SM00430; ZNF_C4_1.
 DR SMART: SM00399; ZNF_C4_1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Steroid-binding.
 FT DOMAIN 1 173 MODULATING.
 FT DNA_BIND 174 239 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 174 194 CA TYPE.
 FT ZN_FING 210 234 CA TYPE.
 FT DOMAIN 240 554 STEROID-BINDING.
 ST SEQUENCE 554 AA; EC13A961CBE89C9 CMC64;
 Query Match 79.78%; Score 1756.5; ID 1; Length 554;
 Best Local Similarity 79.38%; Pred. No. 4.6e-143;
 Matches 330; Conservative 44; Mismatches 49; Indels 3; Gaps 2;
 QY 1 MNYSTPSNVTNLEGGPGKQTTSPNVLWPTPCHSLVVERQLSHLYAFKSTWQKASL 60
 DB 78 MNVNIASNGDSASVROISPSISLWNAHSHSLHQLSILYAFKSTWQKAPPL 137
 QY 61 EHTLVNPKETLKPKKSNQWCAAPV-TGPSKSTAHFVAV-SYVNSVYVWVSPVYKAF 119
 DB 69 EHTLVNPKETLKPKKSNQWCAAPV-TGPSKSTAHFVAV-SYVNSVYVWVSPVYKAF 129
 QY 120 FRSSTLGHVYLTPLNCTTTPKPFSTLALTEPTEVAVNGVNGFEETSELEHPE 179
 DB 129 FRSSTLGHVYLTPLNCTTTPKPFSTLALTEPTEVAVNGVNGFEETSELEHPE 194
 QY 189 KNSFD--CMGKTKYVNAFTVKEILLSTVSPQGVETLLFAEPNVI VSPSPKPTFA 245
 DB 240 SMNMSITKLADKELVHMTSMARKTGFVFLSTFGVRLLESTVMGVLMGIMWSTIDPG 259
 QY 245 SMNMSITKLADKELVHMTSMARKTGFVFLSTFGVRLLESTVMGVLMGIMWSTIDPG 305
 DB 300 KLIFADVLDRDFGKCVGEILEFDMLATTSRELEKLOKREYLVKAMLLNSSKVP 359
 QY 305 KLIFADVLDRDFGKCVGEILEFDMLATTSRELEKLOKREYLVKAMLLNSSKVP 445
 DB 446 LVTATLVANSSPRLAHINAVTDAVWVAWSQTSNQQSMPLMILSLHVGVA 415
 DB 466 LVTATLVANSSPRLAHINAVTDAVWVAWSQTSNQQSMPLMILSLHVGVA 419
 QY 240 SMNMSITKLADKELVHMTSMARKTGFVFLSTFGVRLLESTVMGVLMGIMWSTIDPG 259
 DB 418 SMNMSITKLADKELVHMTSMARKTGFVFLSTFGVRLLESTVMGVLMGIMWSTIDPG 477
 QY 446 KLIFADVLDRDFGKCVGEILEFDMLATTSRELEKLOKREYLVKAMLLNSSKVP 459
 DB 418 KLIFADVLDRDFGKCVGEILEFDMLATTSRELEKLOKREYLVKAMLLNSSKVP 477

Sequence	Best Local Similarity	Prev. No. 1 (40-42)	Matches	Conservative	Mismatches	Indels	Gaps
1. MNVLSPSVNTNLEGGDGEKQITSPNVLMPTGQHSPTLVYVHQLSHL YAEVRSKSPKQKRSR	79, 58	1, 40-42	37	45	6	0	0

CC: Abou04356; BAAT9851.1;
OR HSSB; P03372; UNCO.

InterPro: IP001723; Steroid_receptor.
 InterPro: IP001628; Znf_C4steroid.
 Pfam: PF00104; hormone_rec_1.
 Pfam: PF00105; z1-c4_1.
 PRINTS: PR00398; STEROIDEM-NEP.
 PRINTS: PR00047; STEROIDENGER.
 ProDom: PD00045; Znf_C4steroid_1.
 SMART: SM00430; HOLL1_1.
 SMART: SM00399; Znf_C4_1.
 ProSite: PS00031; NUCLEAR_RECEPTOR; 1.
 KEGG: K01001; Transcription regulation; DNA-binding; Nuclear protein;
 KEGG: K01002; Steroid-binding.
 FT DOMAIN 15 170 MODULATING.
 FT DNA_BIND 171 246 NUCLEAR RECEPTOR-TYPE.
 FT ZN_BINDING 171 191 C4-TYPE.
 FT ZN_BINDING 207 231 C4-TYPE.
 FT ZN_BINDING 237 573 STEROID-BINDING.
 FT DOMAIN 237 573 STEROID-BINDING.
 SO SEQUENCE 573 AA; 63420 MW; 966401D8D39ED4CC CMC64;

Query Match 59.6%; Score 1313.5; DB 1; Length 573;
 Best Local Similarity: 61.4%; Prod. No. 436-105.

Matches 261; Conservative 45; Mismatches 90; Indels 29; Gaps 0.

UY 14 GSP-----CPGTTSPVWVWPT---PCHISPIVWHPQISHLVAEPKSTWCEAFSTEHF 63
 DB 77 GSPAVESPAAHQSLPSLFPMAHGHCHVSPDLHFGQPLVYEPFAHSPWAFKPLEH- 135
 UY 64 LPVNPETLKKKYSKQKSCAPVIGQS-----KRDAHCAVSNVAGHYGVWSEHGT 116
 DB 146 ----GQADTSKLAKRKMAESEPCTSVVATFPAKGMHFAVCHGVASGVHYVWSP-EGC 191
 UY 117 KAFKPSISQGNHNVTPAPNCTDPNPPSPGACPLPCEYEVWVWGSPSPGQVPLV 176
 DB 192 KAFKPSISQGNHNVTPAPNCTDPNPPSPGACPLPCEYEVWVWGSPSPGQVPLV 251
 UY 177 KRSRSD--PQIHWAKKA---KRSQHAHVEHLLALSPQIVLTLLEAPPHV-115 230
 DB 252 RHRKMHTELGLTGAFAFGPFGVAVPTQFAQSALTEQELNREAPPELVM- 311
 UY 231 PPSAPPTVASMMSITKLANKELVHMSWAKHGVFVLSLPGVRLLESGNMVLMGL 293
 DB 312 ELKKPPTESMMMSLNLADKELVHMSWAKHGVFVLSLPGVRLLESGNMVLMGL 371
 UY 291 MMSSTHPPKELFAPDVIPIPEKGVETITFPMILATTSPEPELKAHFFLVYKAM 350
 DB 372 MMSSTHPPKELFAPDVIPIPEKGVETITFPMILATTSPEPELKAHFFLVYKAM 431
 UY 351 ILLNSMYLVATIGDALSSKLAHLNAVITDAVWVIAVSSQSSQNSMTANMIL 400
 DB 432 ILLNSMYLVATIGDALSSKLAHLNAVITDAVWVIAVSSQSSQNSMTANMIL 491
 UY 410 SHVRH 414
 DB 492 AHIRH 496

RESULT 12
 ERM1_CAVAU
 ID ERM1_CAVAU STANDARD: PR1: 568 AA.
 AC G9W669;

DI 16-OCT-2001 (Rel. 40; Created)
 DI 16-OCT-2001 (Rel. 40; Last sequence update)
 DI 15-JUN-2002 (Rel. 41; Last annotation update)
 DI Estrogen receptor beta-1 (Er-beta-1).
 DE Carassius auratus (Goldfish).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprinodontiformes;
 NC Cyprinidae; Carassius.
 OX Ntbl_TaxID=7957;
 RX SEQUENCE FROM N.A.
 RX TISSUE: Liver;
 RX MEDLINE: 99169117; PubMed: 10068500;

Tejoudakova A.V., Pathak S., Callard G.V.;
 "Molecular cloning of an estrogen receptor beta subtype from the
 goldfish, *Carassius auratus*."
 Mol. Cell. Endocrinol. 113:388-400(1999).
 CC FUNCTION: BINDS ESTROGENS WITH AN AFFINITY SIMILAR TO THAT OF
 ER-ALPHA, AND ACTIVATES EXPRESSION OF REPORTER GENES CONTAINING
 ESTROGEN RESPONSE ELEMENTS (ERE) IN AN ESTROGEN DEPENDENT MANNER.
 CC SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 ALPHA (BY SIMILARITY).
 CC SUBCELLULAR LOCATION: Nuclear.
 CC DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.

 This SWISS-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see: <http://www.ebi.ac.uk/stb/charon/charon.html> or send an email to license@sb-eb.ch).

EMBL: AF061269; M226921.1; -.
 DB HSSP: P03372; 1ERR.
 DB InterPro: IP000546; Hormone_rec_1lg.
 DB InterPro: IP001628; Znf_C4steroid.
 DB Pfam: PF00104; hormone_rec_1.
 DB Pfam: PF00105; z1-c4_1.
 DB PRINTS: PR00047; STEROIDENGER.
 DB ProDom: PD00045; Znf_C4steroid_1.
 DB SMART: SM00430; HOLL1_1.
 DB SMART: SM00399; Znf_C4_1.
 DB ProSite: PS00031; NUCLEAR_RECEPTOR; 1.
 KEGG: K01001; Transcription regulation; DNA-binding; Nuclear protein;
 KEGG: K01002; Steroid-binding.
 FT DOMAIN 12 169 MODULATING.
 FT DNA_BIND 170 235 NUCLEAR RECEPTOR-TYPE.
 FT ZN_BINDING 170 190 C4-TYPE.
 FT ZN_BINDING 206 230 C4-TYPE.
 FT ZN_BINDING 236 568 STEROID-BINDING.
 FT DOMAIN 236 568 STEROID-BINDING.
 SO SEQUENCE 568 AA; 63549 MW; 2D67B00775DB76P9 CMC64;

Query Match 56.4%; Score 1243; DB 1; Length 568;
 Best Local Similarity: 59.1%; Prod. No. 550-99;
 Matches 243; Conservative 55; Mismatches 94; Indels 29; Gaps 7;

UY 19 PGTISFNVIWTFCHISPIVWHPQISHLVAEPKSTWCEAFSTEHF 76
 DB 111 PGTISFNVIWTFCHISPIVWHPQISHLVAEPKSTWCEAFSTEHF 131
 UY 93 PGTISFNVIWTFCHISPIVWHPQISHLVAEPKSTWCEAFSTEHF 152
 DB 153 VAETTS TSLGKALMIVGAVCSIVAGSYGVWSTGKAFKRSISQGNHNVTPAPN 210
 UY 77 GNRGAVITGVGSKRIAHGAVCSIVAGSYGVWSTGKAFKRSISQGNHNVTPAPN 136
 DB 153 VAETTS TSLGKALMIVGAVCSIVAGSYGVWSTGKAFKRSISQGNHNVTPAPN 210
 UY 147 GCTIHKPKPSQVAFPEPEYEVWVWGSPSPGQVPLVCEYEVWVWGSPSPGQVPLV 194
 DB 211 GCTIHKPKPSQVAFPEPEYEVWVWGSPSPGQVPLVCEYEVWVWGSPSPGQVPLV 265
 UY 195 SGCHAPVR-----ELIADLSPQIVLTLLEAPPHV-115SRVAPPTESMMMS 244
 DB 267 SGCHAPVR-----ELIADLSPQIVLTLLEAPPHV-115SRVAPPTESMMMS 326
 UY 245 ELKIAPELVHMSWAKHGVFVLSLPGVRLLESGNMVLMGLMMSITKLANKE 304
 DB 326 ELKIAPELVHMSWAKHGVFVLSLPGVRLLESGNMVLMGLMMSITKLANKE 386
 UY 307 TPIAKPFLVHMSWAKHGVFVLSLPGVRLLESGNMVLMGLMMSITKLANKE 366
 DB 387 TPIAKPFLVHMSWAKHGVFVLSLPGVRLLESGNMVLMGLMMSITKLANKE 446
 UY 405 PGLVDPDECKVETITFPMILATTSPEPELKAHFFLVYKAMILNSM-VPLVIA 464
 DB 487 PGLVDPDECKVETITFPMILATTSPEPELKAHFFLVYKAMILNSM-VPLVIA 546
 UY 464 TQVASSPKLAHLNAVITDAVWVIAVSSQSSQNSMTANMILSHVRH 414
 DB 546 TQVASSPKLAHLNAVITDAVWVIAVSSQSSQNSMTANMILSHVRH 600

X MEDLINE: 20218982; PubMed: 10753576;

256 MISWAKIPIGVHLSFTMVHLSQWMEVIMMIMWPSIIPPOKLIIFAPDVLIPPOK

[illegible]

[illegible]

ER HSPD: P01672; 1ERR
 DR InterPro: IPR000536; Hormone_rec_11q.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; Zn_f_4steroid.
 ER Pfam: PF00104; hormone_rec_1.
 DR PRINTS: PR00398; STEROIDRONE.
 DR SMART: SM00430; HOLL: 1.
 DR PROSITE: PS00631; NucleicAcid-binding; NucleicAcid-binding; DNA-binding; Nuclear protein; Zinc-finger; Steroid-binding; Phosphorylation.
 ER RecName: Steroid-binding; Phosphorylation.
 ER Non-ter: 1.
 ER Domain: 1.
 ER ModRes: 251.
 ER Non-ter: 279.
 ER Sequence: 279 AA; 41105 MW; 8580970110A0301 CRO64.
 Query Match: 52.2%; Score 1150; DB 1; Length 279.
 Best Local Similarity: 98.4%; Prod. No. 2.1e-91;
 Matches: 227; Conservative: 2; Mismatches: 2; Indels: 0; Gaps: 0.
 Y 185 QLMARAKAKSGSHARVELLDALSPQLVLTLEAPPHVLSRISAPFTASMMMS 244
 DB 1 QLMARAKAKSGSHARVELLDALSPQLVLTLEAPPHVLSRISAPFTASMMMS 50
 Y 245 LTKLAKELVHMSMAKKIPGVVELSLFQVPLESCWVEVLMGLMPSIDHPKILFA 304
 DB 61 LTKLAKELVHMSMAKKIPGVVELSLFQVPLESCWVEVLMGLMPSIDHPKILFA 120
 Y 305 PDLVLPDEKCEVEGLEIFMLATTSPELKIQHEFYGVKAMILLNSMPPLVAT 364
 DB 121 PDLVLPDEKCEVEGLEIFMLATTSPELKIQHEFYGVKAMILLNSMPPLVAT 180
 Y 365 QANSSPKAHINAVTDAVWVIASSISQSGSMPLANLMLSHVHA 415
 DB 181 QANSSPKAHINAVTDAVWVIASSISQSGSMPLANLMLSHVHA 241
 RESULT 21
 ESRL_CHECK
 ID ESRL_CHECK STANDARD: PRT: 589 AA.
 AC P06212;
 DT 01-JAN-1988 (Ref. 06; Created)
 DT 01-JAN-1988 (Ref. 06; Last sequence update)
 DT 16-OCT-2001 (Ref. 40; Last annotation update)
 DE Estrogen receptor (ER) (estradiol receptor) (ER-alpha).
 GN ESR1 OR NR3A1 OR ESR
 OS Gallus gallus (chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus
 NC NBI_TaxID:9031;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MIMLINE=66247578; PubMed 3755102;
 RA Christ A., Green S., Argus P., Kumar V., Walter P., Bornett J.-M.,
 RA Chabon P.;
 RT "The chicken estrogen receptor sequence: homology with *vertebra* and
 RT the human estrogen and glucocorticoid receptors.";
 RL EMO J. 5:891-897(1986).
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF ENDOCRINE GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC BETA.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
 CC ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
 CC HORMONE-BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
 CC HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA
 CC SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC NR3 SUPERFAMILY.
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/submit/submit.html>).
 CC or send an email to license@ebi.ac.uk.
 CC EMBL: X03805; CA27434.1;
 DR EMBL: S07192; S07192.
 DR HSPD: P01672; 1HCF.
 DR TRANSFAC: T02643;
 DR InterPro: IPR000536; Hormone_rec_11q.
 DR InterPro: IPR001292; oestrgn_receptor.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; Zn_f_4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR PRINTS: PR00398; STEROIDRONE.
 DR PROSITE: PS00631; NucleicAcid-binding; NucleicAcid-binding; DNA-binding; Nuclear protein; Zinc-finger; Steroid-binding.
 DR SMART: SM00399; ZNF_C4; 1.
 DR PROSITE: PS00631; NUCLEAR_RECEPTOR; 1.
 ER RecName: Steroid-binding; Phosphorylation; DNA-binding; Nuclear protein; Zinc-finger; Steroid-binding.
 ER Non-ter: 1.
 ER Domain: 1.
 ER ModRes: 179.
 ER Zn-Fing: 179.
 ER Zn-Fing: 179.
 ER Zn-Fing: 179.
 ER Domain: 245.
 ER Domain: 304.
 ER Domain: 305.
 ER Sequence: 589 AA; 66746 MW; 1B092233C770A54B CRO64.
 Query Match: 48.6%; Score 1070; DB 1; Length 589.
 Best Local Similarity: 51.2%; Prod. No. 4e-84;
 Matches: 214; Conservative: 77; Mismatches: 98; Indels: 28; Gaps: 8;
 Y 22 SPNVLPTRQHSPLVYHR-QLSHVAPQKSPWFAKSTRTLV-----NRPL 71
 DB 100 SPNVLPTRQHSPLVYHR-QLSHVAPQKSPWFAKSTRTLV-----NRPL 155
 Y 72 KPEVSSNPNASPVYGGSSPPDAHPVAVNSDYASGHHVWNSPPKAPPKSLQHHNYI 141
 DB 156 KPEVSSNPNASPVYGGSSPPDAHPVAVNSDYASGHHVWNSPPKAPPKSLQHHNYI 214
 Y 132 GFANVCTIPENPKYQWTELPPEYGVWVFGSPPEPCYLVPPSPSAHPVFWAK 191
 DB 215 GFANVCTIPENPKYQWTELPPEYGVWVFGSPPEPCYLVPPSPSAHPVFWAK 274
 Y 192 AKRSGSHAPV FRTTLD ALSPQLVLTLEAPPHVLS RPSAPFTVA 249
 DB 275 AKRSGSHAPV FRTTLD ALSPQLVLTLEAPPHVLS RPSAPFTVA 334
 Y 249 SPMSLAPADRLVHMSMAKKIPGVVELSLFQVPLESCWVEVLMGLMPSIDHPKILFA 299
 DB 335 SPMSLAPADRLVHMSMAKKIPGVVELSLFQVPLESCWVEVLMGLMPSIDHPKILFA 394
 Y 400 KILFAADLVARDEKCEVEGLEIFMLATTSPELKIQHEFYGVKAMILLNSMPPLVAT 459
 DB 495 KILFAADLVARDEKCEVEGLEIFMLATTSPELKIQHEFYGVKAMILLNSMPPLVAT 554
 Y 460 LVATGTAASSPPAR-ILNAVTDALVWVIASSISQSGSMPLANLMLSHVHA 414
 DB 455 LVATGTAASSPPAR-ILNAVTDALVWVIASSISQSGSMPLANLMLSHVHA 510
 RESULT 22
 ESRL_CHECK
 ID ESRL_CHECK STANDARD: PRT: 587 AA.
 AC P06212;
 DT 01-JAN-1988 (Ref. 06; Created)
 DT 01-JAN-1988 (Ref. 06; Last sequence update)
 DT 16-OCT-2001 (Ref. 40; Last annotation update)
 DE Estrogen receptor (ER) (estradiol receptor) (ER-alpha).
 GN ESR1 OR NR3A1 OR ESR
 OS Gallus gallus (chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus
 NC NBI_TaxID:9031;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MIMLINE=66247578; PubMed 3755102;
 RA Christ A., Green S., Argus P., Kumar V., Walter P., Bornett J.-M.,
 RA Chabon P.;
 RT "The chicken estrogen receptor sequence: homology with *vertebra* and
 RT the human estrogen and glucocorticoid receptors.";
 RL EMO J. 5:891-897(1986).
 CC -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF ENDOCRINE GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC BETA.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
 CC ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
 CC HORMONE-BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
 CC HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA
 CC SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

PA Pellerin H., Fervatelli E., Castagnetta L., Vidal J. J.
KI "Estrogen receptor variant messenger RNA lacking exon 4 in estrogen
EP responsive human breast cancer cell lines.";
PL Cancer Res. 53:741-744(1993).
[4]
RP SEQUENCE OF 110-117, PHOSPHORYLATION, AND MUTAGENESIS.
RX MELINE 96036469; PubMed=7476978;
RA Joel P.H., Traish A.M., Lannigan D.A.;
KI "Estradiol and phorbol ester cause phosphorylation of serine 118 in
PL the human estrogen receptor.";
PL Mol. Endocrinol. 9:1041-1052(1995).
[7]
RP SEQUENCE OF 354-548 FROM N.A.
RC TISSUE-BREAST CARCINOMA;
RA Nandori H., Becker M., Fiebig C., Buettner B., Richter I.;
KI "Mechanisms of acquired tamoxifen resistance in a xenotransplanted
PL human breast carcinoma.";
PL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases
[8]
RP SEQUENCE OF 413-595 FROM N.A.
RA Parker A.;
PL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases
[9]
RP SEQUENCE OF 532-542, AND PHOSPHORYLATION.
RX MELINE 95280953; PubMed=7539106;
RA Arnold S.F., Chouin J.D., Jaffe H., Notides A.C.;
KI "Phosphorylation of the human estrogen receptor on tyrosine 537 in
PL vivo and by src family tyrosine kinases in vitro.";
PL Mol. Endocrinol. 9:24-33(1995).
[10]
RP PHOSPHORYLATION.
RX MELINE 95140025; PubMed=7838153;
RA Arnold S.F., Chouin J.D., Jaffe H., Notides A.C.;
KI "Serine 167 is the major estradiol-induced phosphorylation site on the
PL human estrogen receptor.";
PL Mol. Endocrinol. 8:1208-1214(1994).
[11]
RP PHOSPHORYLATION OF SER-104 AND SER-106, AND MUTAGENESIS.
RX MELINE 95357754; PubMed=10428798;
RA Roulsky L., Trowbridge J.M., Garabedian M.J.;
KI "Potential role of human estrogen receptor alpha transcriptional
PL activation through phosphorylation of serines 104 and 106 by the
PL cys-1 in A-CDK2 complex.";
PL J. Biol. Chem. 274:22296-22302(1999).
[12]
RP STRUCTURE BY NMR OF 180-262.
RX MELINE 91061891; PubMed=2247153;
KA Schwabe J.W.E., Neuhans D., Rhodes D.;
KI "Solution structure of the DNA-binding domain of the estrogen
PL receptor.";
PL Nature 448:458-461(1990).
[13]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 180-262.
RX MELINE 94037103; PubMed=8221895;
KA Schwabe J.W.E., Chapman L., Finch J.T., Rhodes D.;
KI "The crystal structure of the estrogen receptor DNA-binding domain
PL bound to DNA: how receptors discriminate between their response
PL elements.";
PL Cell 75:567-578(1993).
[14]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 305-548.
RX MELINE 97478549; PubMed=9348790;
KA Hrozowski A.M., Pike A.C.W., Dauter Z., Hubbard R.E., Fagan L.;
KI "First structure of estrogen receptor: a dimeric protein with a
PL molecular basis of agonism and antagonism in the estrogen
PL receptor.";
PL Nature 389:753-758(1997).
[15]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 306-544.
RX MELINE 98264297; PubMed=9600906;
KA Jancbaum D.M., Wang Y., Williams S.P., Sigler P.B.;
KI "Crystallographic comparison of the estrogen and progesterone
PL receptor's ligand binding domains.";

PL Proc. Natl. Acad. Sci. U.S.A. 95:5998-6003(1998).
[16]
RP X-RAY CRYSTALLOGRAPHY (1.90 ANGSTROMS) OF 264-554.
RX MELINE 99091051; PubMed=9875847;
RA Shih A.K., Barstad D., Loria P.M., Cheng L., Kushner P.J.;
KI "The structural basis of estrogen receptor/coactivator recognition
PL and the antagonism of this interaction by tamoxifen.";
PL Cell 95:927-937(1998).
[17]
RP 3D-STRUCTURE MODELING OF 411-547.
RX MELINE 98280806; PubMed=9619507;
RA Maalouf G.J., Xu W., Smith T., Mohr S.C.;
KI "Homology model for the ligand-binding domain of the human estrogen
PL receptor.";
PL J. Biomol. Struct. Dyn. 15:841-850(1998).
[18]
RP VARIANT VAL-400.
RX MELINE 9005402; PubMed=2792078;
RA Tora L., Mullick A., Metzger D., Ponglikitmongkol M., Park T.;
RA Chambon P.;
KI "The cloned human estrogen receptor contains a mutation which alters
PL its hormone binding properties.";
PL EMBO J. 8:1981-1986(1989).
[19]
RP VARIANT ALA-447.
RX MELINE 92250550; PubMed=1577818;
KA Reese J.C., Katzenellenbogen B.S.;
KI "Characterization of a temperature-sensitive mutation in the hormone
PL binding domain of the human estrogen receptor. Studies in cell
PL extracts and intact cells and their implications for
PL hormone-dependent transcriptional activation.";
PL J. Biol. Chem. 267:9868-9873(1992).
[20]
RP VARIANT GLU-464.
RX MELINE 97120594; PubMed=8961262;
KA McInerney E.M., Ince B.A., Shapiro D.J., Katzenellenbogen B.S.;
KI "A transcriptionally active estrogen receptor mutant is a novel type
PL of dominant negative inhibitor of estrogen action.";
PL Mol. Endocrinol. 10:1519-1526(1996).
[21]
RP VARIANT CYS-160.
RX MELINE 97346667; PubMed=9195227;
KA Anderson T.L., Wooster R., Laake R., Collins N., Warren W., Skrede M.;
KA Feles R., Twitt K.M., Johnston S.R.D., Dowsett M., Olsen A.O.;
KA Maclellan P., Stratton M.R., Boerresen-Dale A.L.;
KI "Screening for ESR mutations in breast and ovarian cancer patients.";
PL Hum. Mutat. 9:531-536(1997).
[22]
RP -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
PL THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
PL PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
[23]
RP -1- SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
PL BETA.
[24]
RP -1- SUBCELLULAR LOCATION: Nuclear.
[25]
RP -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
PL short form; are produced by alternative splicing.
[26]
RP -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
PL A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
[27]
RP -1- PTM: PHOSPHORYLATED BY CYCLIN A/CDK2. PHOSPHORYLATION PROBABLY
PL ENHANCES TRANSCRIPTIONAL ACTIVITY.
[28]
RP -1- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
PL ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
PL HORMONE BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
PL HORMONE RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISCRETE DNA
PL SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
[29]
RP -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
[30]
RP NR3 SUBFAMILY.
[31]
CC This SWISS PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

This Swiss-prot entry is copyright . It is provided through a collaboration between the Swiss Institute of Bioinformatics and the EMBL station at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed, distorted and/or commercialized; requires a license agreement (<http://www.isb.sib.ch/doc/stand-in-email-to-llicense@isb.sib.ch/>)

08	EMIL2	028954;	RNA25900.1;	
08	EMIL2	A003491;	RNA86925.1;	
08	HSSD	P03372;	IEBR.	
08	InterPro:	IPR000546;	hormone_rece_1fq.	
08	InterPro:	IPR001292;	hormon_recep_1fq.	
08	InterPro:	IPR001743;	steroid_receptor.	
08	InterPro:	IPR001628;	znt_casteroid.	
08	InterPro:	IPR0104;	hormone_rece_1.	
08	Pfam:	PF00105;	z1_c4_1.	
08	Pfam:	PF02159;	horm_recep_1.	
08	PRINTS:	PR00398;	STR000008.	
08	PRINTS:	PR00047;	STR010108.	
08	Prodom:	PD000045;	znt_casteroid_1.	
08	SMART:	SM00400;	W01.1.	
08	SMART:	SM00099;	Z06.04_1.	
08	PROSITE:	PS00041;	N01LEAF_RECEPTOR_1.	
08	Receptor:		Transcription regulation; DNA binding; Nuclear protein;	
08	Receptor:		Transcription binding.	
08	PROPAIR:	1	185	REGULATING.
08	PROPAIR:	186	251	N01LEAF_RECEPTOR_TYPE.
08	PROPAIR:	186	256	C4 TYPE.
08	PROPAIR:	222	246	C4 TYPE.
08	PROPAIR:	252	314	H1R0.
08	PROPAIR:	315	620	STEROID_BINDING.
08	PROPAIR:	249	407	PROLY-GLY.
08	PROPAIR:	320	428	PROLY-LEU.
08	PROPAIR:	511	515	PROLY-LEU.
08	PROPAIR:	576	579	PROLY-GLY.
08	PROPAIR:	620	NA:	6.7729 MW;
08	PROPAIR:	620	NA:	1087018.282kDa5.22 c1064;

[illegible][illegible]

```

RESULT_27
ID      ESTRI_MOUSE    STANDARD          PRT;       599 AA.
A: r1768..5703... 750324;
DEF   01-FEB-1991 (Ref., 17, created)
DEF   01-FEB-1991 (Ref., 17, last sequence update)
DEF   01-JUN-2002 (Ref., 41, last annotation update)
DEF   Estradien receptor (ER) [Estradiol receptor] (Eik-alpha).
DEF   ESR1 or NR1A1 OR ESR OR ESR1 OR ESR1A.
US     Mus musculus (Mouse).
OR     Mammalia: Mammalia: Chordata: Craniata: Vertebrata: Euteleostomi:
OR     Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus:
OX     NCBT TaxId 10090;
[1]
KN      SEQUENCE FROM N.A.
RP      TISSUE: testis;.
RX      MEDLINE=91042558; PubMed=2484714;
RA      White R., Jones J.A., Needham M., Bam J., Parker M.;
RT      "Structural organization and expression of the mouse estrogen
RT      receptor.";
RL      Mol. Endocrinol., 1:735-744(1987);
[2]
KN      SEQUENCE OF 269-599 FROM N.A.
RP      SERIALS 8271 AND 848: S271: Tissue: Spleen;
RA      Ma R.Z., Tenascher C.;
RT      "Screening for candidate genes of mouse autoimmune diseases.";
RL      Submitted (JAN-1999) to the EMBL/GenBank/JDBI databases.
[3]
KN      CARBOHYDRATE LINKAGE SITE SER 575.
RX      MEDLINE=97154020; PubMed=893954;
RA      Jiang M.S., Hart G.W.;
RT      "A superfamily of estrogen receptors are modified by O-linked
RT      N-acetylglucosamine.";
RL      J. Biol. Chem., 272:2421-2426(1997);
[4]
KN      CARBOHYDRATE LINKAGE SITES SER 10; THR 50 AND SER 570.
RX      MEDLINE= 2112487; PubMed=1122684;
RA      Cheng X., Hart G.W.;
RT      "N-glycosylation of the murine estrogen receptor alpha.";
RL      Steroid Biochem. Mol. Biol., 75:147-156(2000);
[5]
DEF   -1- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN THE REGULATION OF ENDOCRINE GENE EXPRESSION AND AFFECT CELLULAR PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
[6]
DEF   + SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERO-DIMER WITH ER ALPHA (BY SIMILARITY).
[7]
DEF   -1- SUBCELLULAR LOCATION: Nucleus.
[8]
DEF   -1- DOMAIN COMPOSED OF THREE DOMAINS: A MODULATING TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
[9]
DEF   + PTM: HYPO-PHOSPHORYLATED BY CYCLIN A/TIRK2 (BY SIMILARITY).
[10]
DEF   -1- MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;

```


[illegible]

Query Match 47.48; Score 1044.5; H1; Length 595;
Best Local Similarity 49.98; Pred. No. 6.2e-82;
Matches 211; Conservative 80; Mismatches 105; Indels 27; Gaps 9

CC	Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Rattus:
CC	NCBI_TaxID: 10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RP	Strain:Wistar;
RA	Morimatsu M.;
EL	Submitted (Mar. 1927) to the EMBL/Genbank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE: 87174780; PubMed:3031601;
RA	Koike S., Sakai M.;
RT	"Molecular cloning and characterization of rat oestrogen receptor
RT	cDNA".;
RL	Nucleic Acids Res. 15:2499-2513(1987).
RN	[3]
RP	SEQUENCE FROM N.A.
RP	Strain:Sprague-Dawley; Tissue:Ovary;
RA	Magel A.M.A.;
RL	Submitted (Jun-1991) to the EMBL/Genbank/DBJ databases.
CC	-1 FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
CC	THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
CC	PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
CC	-1 SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
CC	BETA (BY SIMILARITY).
CC	-1 SUBCELLULAR LOCATION: Nuclear.
CC	-1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC	A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC	-1 FTI: FLUORESCENTLY-LABELLED BY CYCLIN A/CDC2 (BY SIMILARITY).
CC	-1 MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
CC	ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS;
CC	HORMONE-BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
CC	HORMONE-RECEPTOR COMPLEX APPEARS TO RECOGNIZE DISTINCT DNA
CC	SEQUENCES UPSTREAM OF TRANSCRIPTION START SITES.
CC	-1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC	NR3 SUBFAMILY.
CC	-----
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL COLLABORATION
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC	ENTERPRISES REQUIRES A LICENSE FROM SWISS-PROT. WWW.SWISS-PROT.CH
CC	OR SEND AN EMAIL TO LICENSE@SIB.BE.
DR	EMBL: Y00102; CA668287.1;
DR	EMBL: X61098; CA668411.1;
DR	PIR: S07379; S07379.
DR	HSSP: P03772; HCP.
DR	TRANSFAC: T00258;
DE	LOCUS: TFE650536, Bclamc1c1.1g.
DR	InterPro: IP001292; Oestrgn receptor.
DR	InterPro: IP001723; Steroid receptor.
DR	InterPro: IP001628; Znf_C4steroid.
DR	Pfam: PF00104; hormone_recep_1.
DR	Pfam: PF00105; Zf-C4; 1.
DR	Pfam: PF02159; Oest_recept_1.
DR	PRINTS: PR00396; STRHORMONE.
DR	PRINTS: PR00047; STRDIFINER.
DR	Protein: P0000045; Znf_C4steroid; 1.
DR	SMART: SM00430; HOLL; 1.
DR	SMART: SM00359; ZNF_C4; 1.
DR	PROSITE: PS00041; NUCLEAR_RECEPTOR; 1.
KW	Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW	Zinc-finger; Steroid-binding; Phosphorylation.
FT	DOMAIN 1 189 MODULATING;
FT	DNA_BIND 190 255 NUCLEAR RECEPTOR-TYPE.
FT	ZN_FING 190 210 C4-TYPE.
FT	ZN_FING 226 250 C4-TYPE.
FT	DOMAIN 256 315 HINGE.
FT	DOMAIN 316 600 STEROID-BINDING.
FT	DOMAIN 64 71 POLY-ALA.
FT	DOMAIN 171 174 POLY-SER.
FT	MOD_RES 109 109 PHOSPHORYLATION (HY-GWG2) (HY

DR InterPro: IPR000546; Hormone_rec_11q.
 DR InterPro: IPR001292; oestradiol_receptor.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; Zn1_Cassteroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; zf-C4_1.
 DR Pfam: PF02159; oest_recep_1.
 DR PRINTS: PR00398; STERODOM_NER.
 DR PRINTS: PR00477; STERODINERF.
 DR ProDom: PD000045; Zn1_Cassteroid_1.
 DR SMART: SM00430; HOLL_1.
 DR SMART: SM00459; ZNF_C4_1.
 DR PROSITE: PS00031; NUCLEAR_receptor_1.
 KM Receptor: Transcription regulation; DNA-binding; Nuclear protein;
 ZNF11q; Steroid-binding; Alternative splicing.
 FT DOMAIN 1 186 MODULATING.
 FT ZN_FING 187 207 NUCLEAR RECEPTOR-TYPE
 FT ZN_FING 224 247 CA-TYPE.
 FT ZN_FING 254 415 CA-TYPE.
 FT DOMAIN 316 617 HINGE.
 FT DOMAIN 321 425 STEROID-BINDING (POTENTIAL).
 FT DOMAIN 512 516 POLY-LEU.
 FT DOMAIN 577 591 POLY-LEU.
 FT VARSPLIT 1 46 POLY-THR.
 FT VARSPLIT 1 425 MISSING (IN ISOFORM 3).
 SQ SHOOTING 617 AA: 67670 MM: 7820858E5EAD403 CRC64:

Query Match 46.38; Score 1019; DR 1; Length 617;
 Best Local Similarity 49.18; Pred. No. 1e-79;
 Matches 225; Conservative 70; Mismatches 105; Indels 56; Gaps 14;

QY 3 YSI-----DSNVNLESGPSPPTSPNVWPTDGI SPI VNR-----PQLSHYAP- 49
 DB 76 YVAAPDEHNLOPLNG-----SSPPVFPSPSPULSPFIGHPAQGHIAQVVPYLPSS 131
 QY 50 -----QKSPWCEARS-LEHTLPVNRRTLRKV--SNPFAVTPGP-GSKRIARHFAVCS 100
 DB 132 GTSLYRSSVIALSGSVFPLASAPFELVTAVASPSASDPSALHLVRLKRECVS 191
 QY 101 DYASGTHGVGSGEGRKAFKRRSGCHNYGTPANQCTIDNPKPSYQAPVPRGVV 140
 DB 192 DYASGVHYGVSGEGKAFKRRSTQCHNYGVCAINQCTIDNPKRSQACHLRKYEVG 251
 QY 161 MVMGSGRBRGCVRLRPOKSADEQIFKAKKPSGHPR-----VRELL----- 206
 DB 252 MMRGTPRRKTPGPTIKRNPPE-----SGLKFPPTSPAGSSTCVPAIFLQSSG 401
 QY 207 -----LDAISPEQIVLTLEAPPHVLI SRP--SAPTEASMMSLTKIADKFLVMT 257
 DB 405 IEGGATVGVMSPPQVILLIPAEPP-TLQSPQKHERPYSELTINSLINMAHRELVMI 463
 QY 258 SWAKKIPGVEELSLFQVHRLFGQWMEVIMMGLMWSIDHPRKIFAPRIVIDRPRKCV 417
 DB 364 AMAKKVGFDLSLHQVQLTSSWLEFLIMTGLIMHSYTPPKRIFADQLIDKSPGRV 473
 QY 418 PTLLETPMLIATSPPELTKIHKREYLVKAMILLNSMYLVIA-EDGASSSKILAH 476
 DB 424 FCMAPITPDMITAVAPRPIIKIKSPFPVTKAILLNSAFEGSSVPLRPGWVPM 483
 QY 477 LNAVTLALVYAKSSITSSQGSRIANILMLISYVH 414
 DB 484 MONTTALIVYSGSITSVADLSSPEADLTLLSHHFN 721
 RESULT 38
 ESR1_LONGMY STANDARD; PRT; 622 AA.
 ID ESR1_LONGMY
 AC P16058; Q9PSZ9; Q9PSZ8; Q9YGR6;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Estrogen receptor (ER) (estradiol receptor) (ER alpha).

GN ESR1 OR NR4A1 OR ESR.
 GS Oncolytins myxiss (rainbow trout) (Salmo gairdneri).
 GW Enkalypta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 GC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 GZ Flacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 KN [1]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RX MEDLINE=91065824; PubMed=2210031;
 RA Pirkkol E., Le Gall F., Le Gall P., Valotaire Y.;
 RT "Full-length sequence and in vitro expression of rainbow trout
 R1 estrogen receptor cDNA."
 R1 Mol. Cell. Endocrinol. 71:195-204(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
 RD TISSUE-Ovary;
 RX MEDLINE=20114360; PubMed=10650948;
 RA Pirkkol E., Metivier R., Flouriot G., Valotaire Y.;
 RT "Two estrogen receptor (ER) isoforms with different estrogen
 RT dependencies are generated from the trout ER gene."
 RL Endocrinology 141:571-580(2000).
 RN [3]
 RP SEQUENCE OF 150-574 FROM N.A. (SHORT ISOFORM).
 RC TISSUE-Liver;
 RX MEDLINE=8912284; PubMed=2915648;
 RA Pirkkol E., Le Gall F., Valiant G., Le Roux M.G., Valotaire Y.;
 RT "Identification and estrogen induction of two estrogen receptors (ER)
 RT messenger ribonucleic acids in the rainbow trout liver: sequence
 RL homology with other ERS."
 RL Mol. Endocrinol. 3:44-51(1989).
 RN [4]
 RP SEQUENCE OF 270-622 FROM N.A.
 RX MEDLINE=20159854; PubMed=10696781;
 RA Matthews J.B., Zacharewski T.R.;
 RT "Different binding affinities PCRs, HO-PCBs and Aroclors with
 RT recombinant human, rainbow trout (Oncorhynchus mykiss) and green
 RT anole (Anolis carolinensis) estrogen receptors using a semi-high
 RL throughput competitive binding assay."
 RL Toxicol. Sci. 53:426-439(2000).
 CC -1 FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED IN
 CC THE REGULATION OF ENHANCER-GENE EXPRESSION AND AFFECT CELLULAR
 CC PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
 CC -1 SUBUNIT: BINDS DNA AS A HOMODIMER. CAN FORM A HETERODIMER WITH ER-
 CC BETA (BY SIMILARITY).
 CC -1 SUBCELLULAR LOCATION: Nuclear.
 CC -1 ALTERNATIVE PROTEINS: 2 ISOFORMS. A LONG P-PM/ER(1) (SHOWN HERE)
 CC AND A SHORT FORM/ER(5) ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1 DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1 MISCELLANEOUS: IN THE ABSENCE OF LIGAND, STEROID HORMONE RECEPTORS
 CC ARE THOUGHT TO BE WEAKLY ASSOCIATED WITH NUCLEAR COMPONENTS.
 CC HORMONE BINDING GREATLY INCREASES RECEPTOR AFFINITY. THE
 CC HORMONE-RECEPTOR COMPLEX ALWAYS APPEARS TO RECOGNIZE DISTANT DNA
 CC SEQUENCES UPSTREAM OF TRANSCRIPTIONAL START SITES.
 CC -1 SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR3 SUBFAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb.ch/Products/)
 CC or send an email to ilcos@isb.slb.ch).
 CC EMBL: M31559; AAA40552.1;
 DB EMBL: AJ242740; GAB45149.1;
 DB EMBL: AJ242741; GAB45140.1;
 DB EMBL: AF099079; AAD13610.2;
 DB PIR: A37197; A37197;
 DB PIR: A40070; A40070;
 DB HSPD: P04372; IBCO.
 DE InterPro: IPR000546; Hormone_rec_11q.

GenCore version 5.1.6
Copyright (c) 1993 2003 Campogen, LLC

1M protein - protein search, using sw model

Run on: June 16, 2003, 11:34:03 : Search time 23.9719 seconds
(without alignments)
3575, 671 Million cell updates/sec

Title: US-08-826-361a-6

Perfect score: 2203

Sequence: 1 MARYSPSVNTNFKKPSRPT

Q08MR1AM1.M1SHVHPAP 41c

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listed first 120 summaries

Database:

SPRMBL_21:
1: SP archaea:
2: SP bacteria:
3: SP fungi:
4: SP human:
5: SP invertebrate:
6: SP mammal:
7: SP mhc:
8: SP oranelle:
9: SP phage:
10: SP plant:
11: SP rodent:
12: SP virus:
13: SP vertebrate:
14: SP unclassified:
15: SP virus:
16: SP bacteriap:
17: SP archaeap:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length DB	ID	Description
1	2141	97.2	499	6	Q05MFO
2	2058	93.4	486	6	Q05MB9
3	1959	88.9	503	11	Q01786
4	1410.5	64.0	542	11	Q00786
5	1243	56.4	550	6	Q05JCO
6	1241.5	56.4	552	13	Q00MS9
7	1216.5	56.1	553	13	Q00MS9
8	1214.5	56.0	553	13	Q00MS7
9	1220	55.4	565	13	Q00MS7
10	1170.5	53.1	562	13	Q00MS7
11	1065	48.3	587	13	Q00MS7
12	1064	48.3	581	13	Q00MS9
13	1053	47.8	569	13	Q00MS9
14	1046.5	47.5	564	13	Q00MS7
15	1021.5	46.4	441	6	Q05113
16	1017	46.2	578	13	Q00MS7

17	1004.5	45.6	620	13	Q00MS6
18	985.5	44.7	554	13	Q00ZM8
19	961.5	43.6	491	13	Q01033
20	949	43.1	351	13	Q08SM8
21	942.5	42.8	335	13	Q00DZ4
22	924	42.7	276	13	Q01984
23	594.5	27.0	200	13	Q00Z59
24	525	23.8	101	12	Q123V1
25	512	23.2	484	5	Q05SF9
26	512	23.2	406	6	Q08S79
27	506	23.0	249	6	Q05JCI
28	501.5	22.8	710	13	Q01B05
29	476	21.6	583	13	Q01B05
30	468.5	21.5	793	13	Q00VY3
31	460.5	20.9	438	13	Q00ZM7
32	458	20.8	939	6	Q06100
33	457.5	20.8	689	13	Q08GK5
34	457.5	20.8	732	13	Q08D09
35	448	20.3	790	13	P70048
36	442	20.1	87	6	Q00MK2
37	442	20.0	155	13	Q00D63
38	441.5	20.0	401	13	Q06ZM6
39	440.5	20.0	360	13	Q042274
40	437	19.8	688	13	Q00MB8
41	434	19.7	906	4	Q00N21
42	434	19.7	982	6	Q00N08
43	433	19.7	848	13	Q00C09
44	433	19.7	87	6	Q05DK6
45	431.5	19.6	730	4	Q14771
46	430.5	19.5	895	6	Q06K89
47	430	19.5	793	11	Q08530
48	430	19.5	794	11	Q08463
49	430	19.5	866	6	Q06K17
50	429.5	19.5	853	13	Q07245
51	426	19.3	763	13	Q06EV7
52	425	19.3	854	13	Q03244
53	424	19.2	692	13	Q06BP4
54	421.5	19.1	323	4	Q06102
55	421.5	19.1	769	13	Q04497
56	421	19.1	839	13	Q06PV2
57	417.5	19.0	525	6	Q05153
58	417	18.9	563	13	Q00D04
59	416.5	18.9	336	4	Q14267
60	414.5	18.8	340	5	Q05M68
61	413.5	18.8	436	5	Q01506
62	411.5	18.7	303	6	Q07684
63	411.5	18.7	414	11	Q00C00
64	409.5	18.6	400	5	Q01448
65	404.5	18.4	797	13	Q06W65
66	404	18.3	414	5	Q01449
67	403	18.3	763	13	Q00K87
68	402.5	18.3	363	5	Q00AF1
69	401	18.2	344	13	Q01445
70	399.5	18.1	448	5	Q05M63
71	399	18.1	322	11	Q04116
72	392	17.8	412	13	Q01840
73	391.5	17.8	475	5	Q06241
74	389.5	17.7	159	6	Q04402
75	387.5	17.5	427	5	Q00K48
76	381	17.3	448	13	Q00418
77	381	17.3	127	6	Q06RX3
78	380	17.2	379	13	Q00415
79	376	17.1	452	13	Q01613
80	374	17.0	633	6	Q07226
81	373.5	17.0	427	13	Q00D06
82	373	16.9	422	13	Q00417
83	371.5	16.9	434	5	Q05F32
84	371	16.8	348	13	Q01425
85	370.5	16.8	467	5	Q06215
86	370	16.8	480	5	Q05M64
87	370	16.8	489	5	Q007D9
88	367.5	16.7	690	13	Q06T03
89	366.5	16.6	232	13	Q01A30

Q00W16 catenans cat
Q00ZM8 potomyon
Q01033 halothecus
Q08SM8 brachydanto
Q00DZ4 microcerus
Q01984 brachydanto
Q00Z59 brachydanto
Q123V1 meosoteia
Q05SF9 drosophila
Q08S79 drosophila
Q05JCI canis faml
Q01B05 annulla ja
Q01B05 xenopus lae
Q00VY3 xenopus lae
Q00ZM7 potomyon
Q04100 canis faml
Q08GK5 annulla ja
Q08D09 xenopus lae
P70048 xenopus lae
Q00MK2 sus scrofa
Q00D63 micropterus
Q06ZM6 potomyon
Q042274 croceodylus
Q00MB8 oreochromis
Q00N21 homo sapien
Q00N08 salmull sci
Q00C09 annulla ja
Q05DK6 epus cath
Q14771 homo sapien
Q06K89 sus scrofa
Q08530 rattus norv
Q08463 rattus norv
Q06K17 sus scrofa
Q07245 caeca hyacin
Q06EV7 haplochrom
Q03244 ancon hyacin
Q06BP4 haplochrom
Q06102 homo sapien
Q04497 potrus maj
Q06PV2 farassus a
Q05153 mustela vis
Q00D04 halothecus
Q14267 homo sapien
Q05M68 xenopus lae
Q01506 biampylari
Q07684 ovis aries
Q00C00 mus muscul
Q01448 annulla ja
Q01449 annulla ja
Q00K87 oreochromis
Q00AF1 polydora
Q01445 serpens can
Q05M63 atropota mi
Q04116 rattus norv
Q01840 xenopus lae
Q06241 ura pontiat
Q04402 sus scrofa
Q00K48 apus me1111
Q00418 brachydanto
Q06RX3 capra hircu
Q00415 brachydanto
Q01613 xenopus lae
Q07226 sus scrofa
Q00D06 xenopus lae
Q00417 atropota mi
Q05F32 dirot laria
Q01425 xenodophor
Q06215 lucilia cup
Q05M64 atropota mi
Q007D9 leopoda miq
Q06T03 euryzias lat
Q01A30 anolis caro

90	664	16.5	408	5	Q9N9V1	Q9N9V1: Leucocyte beta
91	659	16.3	539	11	Q96JMN	Q96JMN: ratius moty
92	659	16.3	560	11	Q96JMN	Q96JMN: ratius moty
93	659	16.3	627	11	Q96JH6	Q96JH6: mus musculu
94	658	16.3	535	4	Q16413	Q16413: homo sapien
95	655.5	16.1	434	5	P91780	P91780: polyandroa
96	655	16.1	455	11	Q9J319	Q9J319: tamias sibil
97	655	16.1	459	5	Q96S87	Q96S87: aedes aegypt
98	655	16.1	484	5	Q96S88	Q96S88: aedes aegypt
99	654.5	16.1	262	4	Q9H782	Q9H782: homo sapien
100	654	16.1	594	5	Q96041	Q96041: hominid
101	654.5	16.0	125	6	Q19099	Q19099: equus caball
102	650	15.9	538	5	Q77100	Q77100: aedes aegypt
103	650	15.9	565	5	Q77099	Q77099: aedes aegypt
104	649	15.8	661	5	Q26641	Q26641: stramonijace
105	648.5	15.8	485	5	Q904Y3	Q904Y3: aedes albop
106	647	15.8	541	4	Q906C0	Q906C0: homo sapien
107	644	15.6	404	4	Q90B88	Q90B88: homo sapien
108	643.5	15.6	448	14	Q92019	Q92019: equus caball
109	643	15.6	435	5	Q96362	Q96362: triphodacta
110	642	15.5	445	11	Q91841	Q91841: mus musculu
111	641.5	15.5	427	6	Q9T7F0	Q9T7F0: boas taurus
112	641.5	15.5	462	11	P97513	P97513: mus musculu
113	640.5	15.5	731	5	Q9H0P0	Q9H0P0: scistosoma
114	640	15.4	422	14	Q91430	Q91430: brachybanio
115	639.5	15.4	448	14	Q90965	Q90965: brachybanio
116	639	15.4	495	6	Q9GK11	Q9GK11: equus caball
117	638	15.3	411	13	Q9W745	Q9W745: xenopus lae
118	638	15.3	517	13	Q90Y16	Q90Y16: brachybanio
119	637.5	15.3	516	13	Q42186	Q42186: brachybanio
120	636.5	15.3	284	13	Q90Y00	Q90Y00: brachybanio

ALIGNMENTS

RESULT 1

Q95ME0

PRELIMINARY: PRT: 499 AA.

01 DEC-2001 (TREMBLER, 19, created)

01 DEC-2001 (TREMBLER, 19, last sequence update)

01 JUN-2002 (TREMBLER, 21, last annotation update)

ESTROGEN RECEPTOR BETA 2.

ERHETAZ.

Mammalia: Chordata: Craniata: Vertebrata: Euteleostomi:

Mammalia: Euteleostomi: Primates: Catarrhini: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Q9	1	MYSTPSSVNI	60
10	54	MYSTPSSVNI	113
Q9	61	MYSTPSSVNI	120
10	114	MYSTPSSVNI	173
Q9	121	MYSTPSSVNI	180
10	124	MYSTPSSVNI	233
Q9	181	MYSTPSSVNI	240
10	214	MYSTPSSVNI	293
Q9	241	MYSTPSSVNI	300
10	244	MYSTPSSVNI	353
Q9	301	MYSTPSSVNI	360
10	354	MYSTPSSVNI	413
Q9	361	MYSTPSSVNI	416
10	414	MYSTPSSVNI	469

RESULT 2

Q95ME9

PRELIMINARY: PRT: 486 AA.

01 DEC-2001 (TREMBLER, 19, created)

01 DEC-2001 (TREMBLER, 19, last sequence update)

01 JUN-2002 (TREMBLER, 21, last annotation update)

ESTROGEN RECEPTOR BETA 2.

ERHETAZ.

Mammalia: Chordata: Craniata: Vertebrata: Euteleostomi:

Mammalia: Euteleostomi: Primates: Catarrhini: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:

Mammalia: Cercopithecoidea: Cercopithecoidea:


```

DB InterPro: IPR000536: Hormone_rec_114.
DB InterPro: IPR001292: Oestrdn_receptor.
DB InterPro: IPR001724: Stdrnm_receptor.
DB InterPro: IPR001628: ZnI_Castetoid.
DB Pfam: PF00104: hormone_rec_1.
DB Pfam: PF02159: oest_recep_1.
DB Pfam: PF00105: zt-c4_1.
DB PRINTS: PR00398: STROBROMNER.
DB PRINTS: PR00047: STROBDEFNGER.
DB PRODOM: PD000045: ZnI_Castetoid_1.
DB SMART: SM00430: H011_1.
DB SMART: SM00399: Z0F_C4_1.
DB PROSITE: PS00031: NUCLEAR_RECEPTOR; UNKNOWN_1.
KW Receptor.
SQ SOURCE: 587 AA: 66514 MW: 1617288.77 P0513 CRO64.

Query Match: 48.3%; Score 1055; DB 13; Length 587;
Best Local Similarity 52.7%; Pred. No. 4,40-92;
Matches 218; Conservative 69; Mismatches 103; Gaps 24; Gaps 0.

DB 22 SPVAVLTGPHLSPIVVR-QLSHYAFQKSPW--TAPSI FETLPVNFETI PRRVSN 79
DB 100 SPVAVLTGPHLSPIVVR-QLSHYAFQKSPW--TAPSI FETLPVNFETI PRRVSN 156
DB 79 FCAAPVTPG---SKPDHFAVSTAVGVYVWSTPTVAIFKLSGIMV 133
DB 157 FCAAPVTPG---SKPDHFAVSTAVGVYVWSTPTVAIFKLSGIMV 214
DB 134 ATNCTTIDKRRKSGACFELFQVYVWVNSPFPFQVQLVPLVPSAFLQVAVAR 194
DB 217 ATNCTTIDKRRKSGACFELFQVYVWVNSPFPFQVQLVPLVPSAFLQVAVAR 274
DB 194 RSGHAPRV--RELIID-----ALSPQIVLTLEAPRHVILS-PRVAPTFASM 241
DB 277 FAPPTTITWSPVVFIRKYNPA-STAPQMSALFAPPVYVYVYVYVYVYVYVYV 334
DB 242 MMSITKAEKELVHMTSMARKTPGEVLSLPLVPLIFSGVMEVIMMIMMSSTPAPRI 301
DB 437 MLLTNLADELIVHMLNMAVHVGFIATIHGVVHLEFVAMFETIMQVMSGSGSG 396
DB 462 FPAVLTVDGGRKCVYVHLEFVAMFATISPPHLEFVGRKCVYVAVAGVNSGSGV 431
DB 497 LFAVNLIDNGKCVYVHLEFVAMFATISPPHLEFVGRKCVYVAVAGVNSGSGV 554
DB 462 FPAVNLIDNGKCVYVHLEFVAMFATISPPHLEFVGRKCVYVAVAGVNSGSGV 514
DB 457 SSTIKSLERQVIRVIRKITTITITVAKSGSLQVQIRVLAQVLLITSHIRH 510

RESULT 12
Q98SM9 PRELIMINARY: PRT: 581 AA.
AC Q98SM9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Estrogen receptor.
GN ER.
OS Chromidophorus unipareus (Whiptail lizard).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroptossa; Scincomorpha; Teiidae;
OC Teiidae; Chromidophorus.
OX NCBI_TaxID=47197;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 21493797; FJMed 11504222.
KA Sumida K., Oue N., Saito K., Kaneko H.;
KI "Molecular cloning and characterization of reptilian estrogen receptor cDNAs."
KL Mol. Cell. Endocrinol. 184:33-39(2001).
DB EMBL: AB05221; BAB79437.1; -.
DB InterPro: IPR000536: Hormone_rec_114.
DB InterPro: IPR001292: Oestrdn_receptor.

```

```

DB InterPro: IPR001724: Stdrnm_receptor.
DB InterPro: IPR001628: ZnI_Castetoid.
DB Pfam: PF00104: hormone_rec_1.
DB Pfam: PF02159: oest_recep_1.
DB Pfam: PF00105: zt-c4_1.
DB PRINTS: PR00398: STROBROMNER.
DB PRINTS: PR00047: STROBDEFNGER.
DB PRODOM: PD000045: ZnI_Castetoid_1.
DB SMART: SM00430: H011_1.
DB SMART: SM00399: Z0F_C4_1.
DB PROSITE: PS00031: NUCLEAR_RECEPTOR; UNKNOWN_1.
KW Receptor.
SQ SOURCE: 581 AA: 66506 MW: 1697604.44 P03058 CRO64.

Query Match: 48.3%; Score 1064; DB 13; Length 581;
Best Local Similarity 51.0%; Pred. No. 5,40-92;
Matches 214; Conservative 71; Mismatches 95; Gaps 40; Gaps 8.

DB 22 SPVAVLTGPHLSPIVVR-QLSHYAFQKSPW--TAPSI FETLPVNFETI PRRVSN 79
DB 100 SPVAVLTGPHLSPIVVR-QLSHYAFQKSPW--TAPSI FETLPVNFETI PRRVSN 156
DB 79 FCAAPVTPG---SKPDHFAVSTAVGVYVWSTPTVAIFKLSGIMV 133
DB 157 FCAAPVTPG---SKPDHFAVSTAVGVYVWSTPTVAIFKLSGIMV 214
DB 134 ATNCTTIDKRRKSGACFELFQVYVWVNSPFPFQVQLVPLVPSAFLQVAVAR 194
DB 217 ATNCTTIDKRRKSGACFELFQVYVWVNSPFPFQVQLVPLVPSAFLQVAVAR 274
DB 194 RSGHAPRV--RELIID-----ALSPQIVLTLEAPRHVILS-PRVAPTFASM 241
DB 277 FAPPTTITWSPVVFIRKYNPA-STAPQMSALFAPPVYVYVYVYVYVYVYVYV 334
DB 242 MMSITKAEKELVHMTSMARKTPGEVLSLPLVPLIFSGVMEVIMMIMMSSTPAPRI 301
DB 437 MLLTNLADELIVHMLNMAVHVGFIATIHGVVHLEFVAMFETIMQVMSGSGSG 396
DB 462 FPAVLTVDGGRKCVYVHLEFVAMFATISPPHLEFVGRKCVYVAVAGVNSGSGV 431
DB 497 LFAVNLIDNGKCVYVHLEFVAMFATISPPHLEFVGRKCVYVAVAGVNSGSGV 554
DB 462 FPAVNLIDNGKCVYVHLEFVAMFATISPPHLEFVGRKCVYVAVAGVNSGSGV 514
DB 457 SSTIKSLERQVIRVIRKITTITITVAKSGSLQVQIRVLAQVLLITSHIRH 510

RESULT 13
Q98SM9 PRELIMINARY: PRT: 569 AA.
AC Q98SM9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Estrogen receptor alpha.
DE Brachydanio reitio (zebrafish).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID 7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 21493797; FJMed 11504222.
KA Sumida K., Oue N., Saito K., Kaneko H.;
KI "Molecular cloning and characterization of reptilian estrogen receptor cDNAs in zebrafish."
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DB EMBL: AF349412; AKR16740.1; -.
DB HSSP: P03472; IHCP.
DB InterPro: IPR000536: Hormone_rec_114.
DB InterPro: IPR001292: Oestrdn_receptor.
DB InterPro: IPR001724: Stdrnm_receptor.

```


lib 508 LIBRARYCLASSIFICATION 5.27

RESULT 40

LIBRARY PRELIMINARY: PRT: 703 AA.

01-MAR-2002 (TEMBUROL: 20, created)

01-MAR-2002 (TEMBUROL: 20, last sequence update)

01-JUN-2002 (TEMBUROL: 21, last annotation update)

DE: vertebrate receptor (fragment)

OS: Xenopus laevis (African clawed frog)

OC: Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC: Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;

OC: Xenopus; Xenopus;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

RESULT 41

LIBRARY PRELIMINARY: PRT: 438 AA.

01-MAR-2001 (TEMBUROL: 19, created)

01-MAR-2001 (TEMBUROL: 19, last sequence update)

01-JUN-2002 (TEMBUROL: 21, last annotation update)

DE: vertebrate receptor (fragment)

OS: Xenopus laevis (African clawed frog)

OC: Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Hypselostomi;

OC: Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;

OC: Xenopus; Xenopus;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

OC: Nucleotide; Nucleotide;

01-JUN-2002 (TEMBLrel. 21, last annotated update)
 DE Progesterone receptor.
 GN PR.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedata; Canidae; Canis.
 NCBI_TaxID=9615;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Lanthier-van Leuven J S., van Garderen E., Mol J A.;
 RT "Molecular cloning and cellular localization of the canine
 progesterone receptor."
 RT Submitted (Aug-1999) to the EMBL/Genbank/DDBP databases
 CC -1- SUPPLEMENTARY LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 DR EMBL: AF177470; A069282.1;
 DR HSSP: P06401; 1A28;
 DR InterPro: IPR000346; Hormone_Recep_1;
 DR InterPro: IPR001283; Progesterone_Receptor;
 DR InterPro: IPR001723; Steroid_Receptor;
 DR InterPro: IPR001628; Zinc_Castrolid;
 DR Pfam: PF00104; hormone_recep_1;
 DR Pfam: PF02161; Progesterone_Receptor_1;
 DR Pfam: PF00185; ZINC_1;
 DR PRINTS: PR00398; STERHORMONER.
 DR PROSITE: PROSITE7: STER-RECEPTOR.
 DR PRODOM: PD000035; Zinc_Castrolid; 1.
 DR SMART: SM00399; ZINC_1;
 DR SMART: SM00399; ZINC_1;
 DR PROSITE: PROSITE7: NUCLEAR_RECEPTOR_1;
 DR DNA-binding: Nuclear protein: Receptor; Transcription regulation;
 KW Zinc-finger.
 SQ SEQUENCE 939 AA: 98417 MW: 144863564239 GRC64.

Query Match 20.8% Score 458; DB 6; Length 939;
 Best Local Similarity 32.5%; Prod No 2674;
 Matches 106; Conservative 64; Mismatches 126; Indels 43; Gaps 10;

96 CAVCTSYASVYHGVWSTEGCKAFKRSIOGHNYICPATNCTIDKPKKSCQAQPLK 155
 573 CILCGFASCHYGVITGSGVTFKRAKRGHNTVAPNITVTKIPKNTACTPIK 632
 DB 156 CEEGVNWKSGSPREKQGVLPVPRQPSAEQD--HCAKAKPSGSHAPVVKLLIALSPK 213
 DB 633 CCGACGVLAGPKFK--FNKRVVMTLLDAVALPQVGIPIESQALSGIT-----SEPS 684
 QY 214 Q-----LVTLLEKPPHVLISRSA-PTEASMMMSLTKLADKELVIMISWAKKIFG 265
 DB 685 QDILGLPLINILMSLPIVIVAGHINTKPIPTSSLSLSLNDJDEKQLSVYKWSKSLPG 744
 QY 266 PVEELSTPQVPLESCWMEVLMCMGMSSTHFG--KLIFACVADPKKRVWGLLEI 323
 DB 745 FRLIATIDQITLIGVWMSLWFGVGSYKRVASQIMYPAHILINLQPKK--ESVPSI 803
 QY 324 FFMILATTSPPPELKTQKEVYGVKAMILLNSMTGVTATVPAISSEYLAHLIAVIVA 383
 DB 804 CLTMMQIPQGFVKIYVSGPEFLVMTLLINT--PL-----HGLPSQWTFEMPSYIRE 857
 QY 384 LVMVIA--KSGISSQOOSMLANL 406
 DB 858 LIRATGLRQKGVSSQSPFYHITKIL 883

RESULT 33
 Q98GX5 PRELIMINARY: PRT: 689 AA.
 ID Q98GX5
 AC Q98GX5
 DT 01-JUN-2002 (TEMBLrel. 21, Created)
 DT 01-JUN-2002 (TEMBLrel. 21, last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, last annotation update)
 DE Progesterone receptor type 2
 OS Anguilla japonica (Japanese eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Anguilla japonica. Anguilla japonica; Anguilla japonica;
 OC Anguilla.
 GN NCBI_TaxID=7937;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=TESTIS;
 KX MEDLINE=21629085; PubMed=11755535;
 RA Ikemuchi T., Todo T., Kobayashi T., Nagahama Y.;
 RT "A novel progesterone receptor subtype in the Japanese eel, Anguilla
 japonica."
 RT FEBS Lett. 510:77 82(2002).
 DR EMBL: AB028024; BAB85993.1;
 KW Receptor
 SQ SEQUENCE 689 AA: 76316 MW: 1401650672726 GRC64.

Query Match 20.8% Score 457.5; DB 13; Length 689;
 Best Local Similarity 30.3%; Prod No. 15634;
 Matches 118; Conservative 72; Mismatches 157; Indels 43; Gaps 12;

41 QLSHYAEQKSPWCEAPSELETLIPNPELTLPKVSQNPYA-----SPVTSPSSKPP 92
 DB 263 KSLPTSPSSQSWDSTGLSEDFEFTGYLPGLIPNICTHNSLKSQVYMGMSQK- 321
 QY 93 AHPQAVOSVYASVYHGVWSTEGCKAFKRSIOGHNYICPATNCTIDKPKKSCQAQPLK 152
 DB 322 CILCGFASCHYGVITGSGVTFKRAKRGHNTVAPNITVTKIPKNTACTPIK 379
 QY 153 CEEGVNWKSGSPREKQGVLPVPRQPSAEQD--HCAKAKPSGSHAPVVKLLIALSPK 203
 DB 380 LIRQGVAGVMTIGSPKMKKLSALKVLGLDQSLAVRSPICASYF-----GQALAILSPKVR 435
 QY 204 ELTDLSPQVLETLLEKPPHVLISRSA-PTEASMMMSLTKLADKELVIMISWAK 261
 DB 436 EL---QFTPO--ILSTLENIPEPTYSGVDATQETPELLNSLGLFRLIMIVKRSK 490
 QY 262 KIPGEVPSIPQVPLESCWMEVLMCMGMSSTHFG--KLIFACVADPKKRVWGLLEI 319
 DB 491 SLGPFSLHINDQMTLIQYWSMSLWFGVGSYKRVASQIMYPAHILINLQPKK----- 546
 QY 320 LLELEMLLA LSKFELKELKKEVLYVAMILLNSMTGVTATVPAISSEYLAHLIAVIVA 376
 DB 547 NSPTIDCMAMQIIPQAFNLQVYKEEFLQKMLVLLINT--VPLEGRLSQAPDPRKHGY 604
 QY 377 INAVTDAIWMVIAKSGISSQOOSMLANL 406
 DB 605 IPIETKA-IQITPEGVASSQPFYHITKIM 633

RESULT 34
 Q9DD09 PRELIMINARY: PRT: 732 AA
 ID Q9DD09
 AC Q9DD09
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, last annotation update)
 DE Progesterone receptor.
 GN XPR-1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20570484; PubMed=11114187;
 RA Tian J., Kim S., Helling F., Ruderman J.V.;
 RT "Identification of XPR-1, a progesterone receptor required for Xenopus
 oocyte activation."
 RT Proc. Natl. Acad. Sci. U S A. 97:14358-14364(2000).
 CC -1- SUPPLEMENTARY LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 DR EMBL: AF279335; AN042362.1;
 DR HSSP: P06536; 1R6D.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleotide search, using sw model

Run on: June 16, 2003, 00:23:52 ; Search time 3329.11 seconds

(without alignments)
10988 581 Million cell updates/sec

Title: US-08-826-361A-20

Percent score: 1257

Sequence: 1 ATGATTAACACATTCCTCA

US-08-826-361A-20

Scoring table:

IDENTITY_MDC

Gapop 10.0 ; Gapext 1.0

Searched:

2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:

4109280

Minimum hit seq length: 0
Maximum hit seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 120 summaries

Data base:

1: gb_ba: *
2: gb_ba: *
3: gb_ba: *
4: gb_ba: *
5: gb_ba: *
6: gb_ba: *
7: gb_ba: *
8: gb_ba: *
9: gb_ba: *
10: gb_ba: *
11: gb_ba: *
12: gb_ba: *
13: gb_ba: *
14: gb_ba: *
15: gb_ba: *
16: gb_ba: *
17: gb_ba: *
18: gb_ba: *
19: gb_ba: *
20: gb_ba: *
21: gb_ba: *
22: gb_ba: *
23: gb_ba: *
24: gb_ba: *
25: gb_ba: *
26: gb_ba: *
27: gb_ba: *
28: gb_ba: *
29: gb_ba: *
30: gb_ba: *
31: gb_ba: *
32: gb_ba: *
33: gb_ba: *
34: gb_ba: *
35: gb_ba: *
36: gb_ba: *
37: gb_ba: *
38: gb_ba: *
39: gb_ba: *
40: gb_ba: *
41: gb_ba: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1257	100.0	1257	6	AF5124	AF5124 Sequence 20
2	1257	100.0	1257	6	AX323474	AX323474 Sequence
3	1250	99.4	99.4	5	AF060555	AF060555 Homo sapi
4	1249	99.4	1251	6	AF5105	AF5105 Sequence 2
5	1249	99.4	1251	6	AX323056	AX323056 Sequence
6	1249	99.4	1251	6	AF051428	AF051428 Homo sapi
7	1247	99.2	1247	6	AF060589	AF060589 Homo sapi
8	1247	99.2	1247	6	AF5104	AF5104 Sequence 1
9	1247	99.2	1247	6	AX323055	AX323055 Sequence
10	1247	99.2	1247	6	AF1586	AF1586 Sequence 4
11	1247	99.2	1247	6	AR075921	AR075921 Sequence
12	1247	99.2	1247	6	AX023458	AX023458 Sequence
13	1247	99.2	1247	6	HSBNAERB	HSBNAERB
14	1247	99.2	1247	6	AX029400	AX029400 Sequence
15	1247	99.2	1247	6	AB006590	AB006590 Homo sapi
16	1247	99.2	1247	6	AF5127	AF5127 Sequence 24
17	1247	99.2	1247	6	AX323078	AX323078 Sequence
18	1247	99.2	1247	6	AX323458	AX323458 Sequence
19	1247	99.2	1247	6	AF051427	AF051427 Homo sapi
20	1186.4	94.4	1640	9	AF394815	AF394815 Macaca ar
21	1127.2	89.7	1566	9	AF394816	AF394816 Gallitrit
22	1125.4	89.5	1651	9	CJEDRB	CJEDRB
23	995	79.2	1584	4	AF177936	AF177936 oviss ario
24	987	78.5	1611	4	AF110402	AF110402 Bos tauru
25	985.4	78.4	1611	4	AF110402	AF110402 Bos tauru
26	963	76.6	1581	4	AF164957	AF164957 Sus scrofa
27	963	76.6	1581	4	AF164957	AF164957 Sus scrofa
28	960.8	76.4	1427	9	AF124790	AF124790 Homo sapi
29	960.8	76.4	1427	9	BC024181	BC024181 Homo sapi
30	940.2	74.8	1421	10	AF042058	AF042058 Rattus no
31	940.2	74.8	1421	10	AF042058	AF042058 Rattus no
32	937	74.5	1585	10	AF074436	AF074436 Rattus no
33	937	74.5	1585	10	AF074436	AF074436 Rattus no
34	937	74.5	1585	10	AF074436	AF074436 Rattus no
35	937	74.5	1585	10	AF074436	AF074436 Rattus no
36	929	73.9	1512	10	MM081451	MM081451 Mus musculu
37	922.6	73.4	1458	6	AF075922	AF075922 Sequence
38	922.6	73.4	1458	6	AF075922	AF075922 Sequence
39	922.6	73.4	1458	6	AX022460	AX022460 Sequence
40	922.6	73.4	1458	6	AX022460	AX022460 Sequence
41	876.2	69.7	1512	10	AF042059	AF042059 Rattus no
42	876.2	69.7	1512	10	AF042059	AF042059 Rattus no
43	876.2	69.7	1512	10	AF042059	AF042059 Rattus no
44	865	68.8	1512	10	AF054413	AF054413 Mus muscu
45	865	68.8	1512	10	AF054413	AF054413 Mus muscu
46	865	68.8	1512	10	AF054413	AF054413 Mus muscu
47	865	68.8	1512	10	AF054413	AF054413 Mus muscu
48	865	68.8	1512	10	AF054413	AF054413 Mus muscu
49	865	68.8	1512	10	AF054413	AF054413 Mus muscu
50	865	68.8	1512	10	AF054413	AF054413 Mus muscu
51	865	68.8	1512	10	AF054413	AF054413 Mus muscu
52	865	68.8	1512	10	AF054413	AF054413 Mus muscu
53	865	68.8	1512	10	AF054413	AF054413 Mus muscu
54	865	68.8	1512	10	AF054413	AF054413 Mus muscu
55	865	68.8	1512	10	AF054413	AF054413 Mus muscu
56	865	68.8	1512	10	AF054413	AF054413 Mus muscu
57	865	68.8	1512	10	AF054413	AF054413 Mus muscu
58	865	68.8	1512	10	AF054413	AF054413 Mus muscu
59	865	68.8	1512	10	AF054413	AF054413 Mus muscu
60	865	68.8	1512	10	AF054413	AF054413 Mus muscu
61	865	68.8	1512	10	AF054413	AF054413 Mus muscu
62	865	68.8	1512	10	AF054413	AF054413 Mus muscu
63	865	68.8	1512	10	AF054413	AF054413 Mus muscu
64	865	68.8	1512	10	AF054413	AF054413 Mus muscu
65	865	68.8	1512	10	AF054413	AF054413 Mus muscu

pred. No. is the number of results predicted by chance to have 1

Db 641 TGTGATGGAGGCTGTTAATGATGAGGAGTGGAGTGGATTTGATGATGATGAGGAG 900
 QY 901 CTGATCTTTGCTTCCAGATCTTCTTCTGACACAGCATGAGGAGAAATGATGAGAAAT 960
 Db 901 CTGATCTTTGCTTCCAGATCTTCTTCTGACACAGCATGAGGAGAAATGATGAGAAAT 960
 QY 961 CTGAGAACTTGGATGATGCTGGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1020
 Db 961 CTGAGAACTTGGATGATGCTGGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1020
 QY 1021 CAAAGAAATATCTGCTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1080
 Db 1021 CAAAGAAATATCTGCTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1080
 QY 1081 GTGACAGGAGGAGATCTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1140
 Db 1081 GTGACAGGAGGAGATCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1140
 QY 1141 AGGCAATCTTGGATGATGCTGGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1200
 Db 1141 AGGCAATCTTGGATGATGCTGGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1200
 QY 1201 GAGTGGATTAATCTTCTGATGCTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1260
 Db 1201 GAGTGGATTAATCTTCTGATGCTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1260

RESULT 2
 AX123074 1257 bp DNA Linear PAT 07 JAN 2002
 DEFINITION Sequence 20 from patent EP1162264.
 AX123074
 VERSION AX123074.1 GI:18093957
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

REFERENCE
 AUTHORS Mosselman, S. and Dijkema, R.
 TITLE ChimERIC hormone receptor
 JOURNAL Patent: EP 1162264 A 20 12-DEC 2001.
 Akzo Nobel N.V. (NL)
 FEATURES
 source 1..1257
 location/junctions
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 297 a 327 c 355 g 278 t
 ORIGIN

Query Match 100.0% Score 1257; DB 5; Length 1257;
 Best local similarity 100.0%; Pred. No. 44-274;
 Matches 1257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGATTTACAGCATTTGAGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 60
 Db 1 ATGATTTACAGCATTTGAGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 60
 QY 61 ACAAGGCAATATGCTTGGGCAAGATCTTCTGAGGCTTGGAGATTTAAATTTCAA 120
 Db 61 ACAAGGCAATATGCTTGGGCAAGATCTTCTGAGGCTTGGAGATTTAAATTTCAA 120
 QY 121 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 180
 Db 121 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 180
 QY 181 GAAATACGCTTACGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 240
 Db 181 GAAATACGCTTACGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 240
 QY 241 GGTAGGCTTACTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 300
 Db 241 GGTAGGCTTACTGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 300

QY 401 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 460
 Db 401 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 460
 QY 461 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 520
 Db 461 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 520
 QY 521 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 580
 Db 521 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 580
 QY 581 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 640
 Db 581 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 640
 QY 641 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 700
 Db 641 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 700
 QY 701 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 760
 Db 701 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 760
 QY 761 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 820
 Db 761 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 820
 QY 821 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 880
 Db 821 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 880
 QY 881 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 940
 Db 881 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 940
 QY 941 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1000
 Db 941 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1000
 QY 1001 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1060
 Db 1001 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1060
 QY 1061 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1120
 Db 1061 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1120
 QY 1121 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1180
 Db 1121 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1180
 QY 1181 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1240
 Db 1181 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1240
 QY 1241 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1300
 Db 1241 CATTATTCATCTGATGCTGAGGCAATCTTCTGAGGCTTGGAGATTTAAATTTCAA 1300

RESULT 4
 AF060555 2745 bp mRNA Linear PAT 29 APR 1998
 DEFINITION Homo sapiens estrogen receptor beta 3 isoform mRNA, complete cds.
 AF060555
 ACCESSION AF060555.1 GI:3091285
 VERSION
 KEYWORDS

SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2745)
AUTHORS Moore, J. T., McKee, J. D., Slonitz-Kosler, K., Moore, L. B., Jones, S. A., Su, J. L., Horton, E. L., Kiew, S. A., Lachmann, J. M., and Wilson, F. M.
TITLE Direct Submission
JOURNAL Submitted (20 APR 1998) Molecular Sciences, Glaxo Wellcome, 5 Moore Drive, RTP, NC 27709, USA

FEATURES
SOURCE
1..2745
/organism="Homo sapiens"
/db_xref="taxon:9606"
471..2012
/note="splice variant"
/product="estrogen receptor beta 4 isoform"
/protein_id="AF15234.1"
/db_xref="GI:4091286"
/translation="MOTKNSPSI NSPSYNSQSLLPIPIHSIVTSSVSHSEVP
AMFVSAVMNYSIPSNVNLKGGKQDTSINVLPTSHLSPLVHOLSHYAP
OKSPWEARSLKEFLPNEETLAKRVSINRQASVYDPSKRIAHFAVPSDASGH
YGVMSGEKFEKPKS LGCHNY LGPAINQTLIDKREKSGQAKRKYVQAKVQ
SPREKRYELVPHQSAIRQILAKAKRSQSHAPREELDLALSPQVLETLAE
PQVILSRSAVPEASMMKSTIKLADKELVHISWAKIPEFVELSLDQVLESL
WMEVLMKIMWSIDHPCKELFAPILVILDEKQVGELEIFIMLLATSRFRELK
QIREVLYKAMILLNSMSPVLTATVDIMSSPFLMLINAVETALVWVLAQSLSSQ
QSRKLANLMLLSHVMHSSLSIMRLMLKFAKSHVQKDTQVCAHMSVSRSEFAN
QDREK"

BASE COUNT 741 a 665 c 690 g 649 t
ORIGIN

Query Match 99.98% Score 1250.6; DB 9; Length 2745;
Best Local Similarity 99.78% Proc. No. 2,562,272;
Matches 1254; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 ATGATTTATACATCTGCATCTCAATCTCAATCTTGGAAGGTGGAGCTGGTGGACACG 60
|||||
640 ATGATTTATACATCTGCATCTCAATCTCAATCTTGGAAGGTGGAGCTGGTGGACACG 60
|||||
61 ACAAATCT 120
|||||
690 ACAAATCT 749
|||||
121 CACTATCATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCT 180
|||||
750 CACTATCATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCT 809
|||||
181 GAATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 240
|||||
810 GAATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 869
|||||
241 GATATCT 300
|||||
870 GATATCT 929
|||||
401 GATTATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 460
|||||
940 GATTATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 989
|||||
961 AATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 420
|||||
990 AATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1049
|||||
421 GATATATCT 480
|||||
1050 GATATATCT 1109
|||||
481 ATATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 540
|||||
1110 ATATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1169
|||||

541 AGTGGAGACACACCTGACCTGCTGCTGGACAGAGAGACACACACACACACACACAC 600
|||||
1170 AGTGGAGACACACCTGACCTGCTGCTGGACAGAGAGACACACACACACACACACAC 1229
|||||
601 GATATCT 660
|||||
1240 GATATCT 1289
|||||
661 GATATCT 720
|||||
1290 GATATCT 1349
|||||
721 ATATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 780
|||||
1350 ATATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1409
|||||
781 AATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 840
|||||
1410 AATATCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCT 1469
|||||
841 TCTTCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
|||||
1470 TCTTCATCTCTATCTCTCAATCTCAAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1529
|||||
901 CTATATCT 960
|||||
1530 CTATATCT 1589
|||||
961 CTATATCT 1020
|||||
1590 CTATATCT 1649
|||||
1021 CACAAATATCTCTCTCTCAAGACATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
|||||
1650 CACAAATATCTCTCTCTCTCAAGACATATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1709
|||||
1081 CTATATCT 1140
|||||
1710 CTATATCT 1769
|||||
1141 AATATCT 1200
|||||
1770 AATATCT 1829
|||||
1201 GATATCT 1257
|||||
1830 GATATCT 1886
|||||

RESULT 4
AB5105 1251 bp RNA linear PAT 29 MAR 1999
DEFINITION Sequence 2 from Patient BP0798478.
ACCESSION AB5105
VERSION AB5105.1 GI:4530974
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Mosselman, S., and Dijkema, R.
TITLE Estrogen receptor
JOURNAL Patient: BP 0798378-A 2 01-06T-1997;
COMMENT AKZO Model NV (NL)
ORIG publication CA 2290423 19970926
ORIG publication AU 1652197 19971002.
FEATURES
SOURCE
1..1251
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 297 a 324 c 454 g 276 t
ORIGIN

Query Match: 99.4%; Score 1249; DB 6; Length 1251;
 Best Local Similarity 100.0%; Pred. No. 5,4e-272;
 Matches 1249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ATGAAATTACAGCATTCCTGAGCAATCTCACTTGAAGCTGGAGCTGGTGGAGAGAG 60
 1 ATGAAATTACAGCATTCCTGAGCAATCTCACTTGAAGCTGGAGCTGGTGGAGAGAG 60
 Db 1 ATGAAATTACAGCATTCCTGAGCAATCTCACTTGAAGCTGGAGCTGGTGGAGAGAG 60

Y 61 A AAGGCAATATGTTGTTGGCAACACTGGGCAAGCTTTCTCTTAGTGGTCAAGC 120
 61 A AAGGCAATATGTTGTTGGCAACACTGGGCAAGCTTTCTCTTAGTGGTCAAGC 120
 Db 61 A AAGGCAATATGTTGTTGGCAACACTGGGCAAGCTTTCTCTTAGTGGTCAAGC 120

Y 121 CAGTATACATCTGATGCGGACCTGAAAAAGATGCTGGTGGAGCAAGATGGCTA 180
 121 CAGTATACATCTGATGCGGACCTGAAAAAGATGCTGGTGGAGCAAGATGGCTA 180
 Db 121 CAGTATACATCTGATGCGGACCTGAAAAAGATGCTGGTGGAGCAAGATGGCTA 180

Y 181 GAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 240
 181 GAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 240
 Db 181 GAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 240

Y 241 GCAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 300
 241 GCAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 300
 Db 241 GCAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 300

Y 301 GATTACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 360
 301 GATTACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 360
 Db 301 GATTACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 360

Y 361 AAAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 420
 361 AAAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 420
 Db 361 AAAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 420

Y 421 GATAAAACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 480
 421 GATAAAACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 480
 Db 421 GATAAAACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 480

Y 481 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 540
 481 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 540
 Db 481 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 540

Y 541 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 600
 541 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 600
 Db 541 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 600

Y 601 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 660
 601 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 660
 Db 601 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 660

Y 661 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 720
 661 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 720
 Db 661 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 720

Y 721 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 780
 721 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 780
 Db 721 ATATTAATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 780

Y 781 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 840
 781 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 840
 Db 781 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 840

Y 841 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 900
 841 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 900
 Db 841 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 900

Y 901 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 960
 901 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 960
 Db 901 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 960

Y 961 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1020
 961 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1020
 Db 961 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1020

Y 1021 GATTAAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1080
 1021 GATTAAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1080
 Db 1021 GATTAAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1080

Db 1021 GATTAAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1080
 1021 GATTAAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1080
 Db 1021 GATTAAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1080

Y 1081 GTCACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGG 1140
 1081 GTCACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGG 1140
 Db 1081 GTCACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGG 1140

Y 1141 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1200
 1141 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1200
 Db 1141 AAGAGATCTGCGCTTCTGAGAGCTCAAGCTGTTGAGCAAGTGGTGGAGCAAG 1200

Y 1201 GTCACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGG 1260
 1201 GTCACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGG 1260
 Db 1201 GTCACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGGACAGG 1260

RESULT 5
 AX323056 1251 bp LMA 110001 PM 07-JAN-2002
 DEFINITION
 AX323056
 VERSION
 AX323056.1 GI:18094943
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
 1
 Messelman, S. and D.J. Kema, P.
 TITLE
 Chimeric hormone receptor
 Patent: EP 1162264-A 2 12 DEC 2001;
 JOURNAL
 Akzo Nobel N.V. (NL)

FEATURES
 SOURCE
 1..1251
 Location/Qualifiers
 /db_xref="taxon,9606"

FAST LENGTH 297 a 424 c 454 g 276 t
 ORIGIN

Query Match: 99.4%; Score 1249; DB 6; Length 1251;
 Best Local Similarity 100.0%; Pred. No. 5,4e-272;
 Matches 1249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 ATGAAATTACAGCATTCCTGAGCAATCTCACTTGAAGCTGGAGCTGGTGGAGAGAG 60
 1 ATGAAATTACAGCATTCCTGAGCAATCTCACTTGAAGCTGGAGCTGGTGGAGAGAG 60
 Db 1 ATGAAATTACAGCATTCCTGAGCAATCTCACTTGAAGCTGGAGCTGGTGGAGAGAG 60

Y 61 A AAGGCAATATGTTGTTGGCAACACTGGGCAAGCTTTCTCTTAGTGGTCAAGC 120
 61 A AAGGCAATATGTTGTTGGCAACACTGGGCAAGCTTTCTCTTAGTGGTCAAGC 120
 Db 61 A AAGGCAATATGTTGTTGGCAACACTGGGCAAGCTTTCTCTTAGTGGTCAAGC 120

Y 121 CAGTATACATCTGATGCGGACCTGAAAAAGATGCTGGTGGAGCAAGATGGCTA 180
 121 CAGTATACATCTGATGCGGACCTGAAAAAGATGCTGGTGGAGCAAGATGGCTA 180
 Db 121 CAGTATACATCTGATGCGGACCTGAAAAAGATGCTGGTGGAGCAAGATGGCTA 180

Y 181 GAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 240
 181 GAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 240
 Db 181 GAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 240

Y 241 GCAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 300
 241 GCAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 300
 Db 241 GCAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 300

Y 301 GATTACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 360
 301 GATTACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 360
 Db 301 GATTACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 360

Y 361 AAAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 420
 361 AAAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 420
 Db 361 AAAACACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 420

Y 421 GATAAAACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 480
 421 GATAAAACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 480
 Db 421 GATAAAACATCTACCTCTAAATAGAGACATCTGAAAAAGATGCTGGTGGAGCAAG 480

UY 781 AACAAATTCCTGCTTTCTGAGCTCAACCTCTTCTGAACTCAAGCTCTCTGAGAGC 840
|||||
1b 2215 AACAAATTCCTGCTTTCTGAGCTCAACCTCTTCTGAACTCAAGCTCTCTGAGAGC 2274
UY 841 TCTTCAATGAGAGCTTCAATGAGAGCTGATGCTGCTTCAATGAGAGCTGAGAGC 900
|||||
1b 2275 TCTTCAATGAGAGCTTCAATGAGAGCTGATGCTGCTTCAATGAGAGCTGAGAGC 2344
UY 901 CTATCTCTCTGCTCAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 960
|||||
1b 2345 CTATCTCTCTGCTCAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAT 2394
UY 961 CTGCAAAATCTTGAATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
|||||
1b 2395 CTGCAAAATCTTGAATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2454
UY 1021 CAAAGAAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
|||||
1b 2455 CAAAGAAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2514
UY 1081 GTTCAAG 1140
|||||
1b 2515 GTTCAAG 2574
UY 1141 ATTCAGCTCTTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
|||||
1b 2575 ATTCAGCTCTTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2634
UY 1201 GCTCTCTCTGAG 1252
|||||
1b 2635 GCTCTCTCTGAG 2696

RESULT 8
A65104 1444 bp DNA linear PAT 29-MAR-1999
DEFINITION
A65104
ACCESSION
A65104.1 GI:4540973
KEYWORDS
SOURCE
organism
unidentified
unclassified
REFERENCE
1. (bases 1 to 1444)
TITLE
Estrogen receptor
PATENT: EP 0798478 A1 01-09-1997;
JOURNAL: AKZO NOBEL NV (NL)
COMMENT
of her publication CA 2200423 19970926
of her publication AU 1652197 19971002
FEATURES
SOURCE
1. 1444
/organism "unidentified"
/db_xref "taxon:32644"

BASE COUNT 447 a 373 c 406 g 308 t
GRTSIN

Query Match 99.2% Score 1247; DB 6; Length 1444;
Exact Local Similarity 100.0%; Prev. No. 1,66; 271;
Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 181 GAAGAAATCTTACCTGAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
|||||
1b 181 GAAGAAATCTTACCTGAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
UY 241 GTCAAGCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
|||||
1b 241 GTCAAGCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
UY 361 GATTAAG 360
|||||
1b 401 GATTAAG 360
UY 461 AAAAG 420
|||||
1b 461 AAAAG 420
UY 481 GATTAAG 480
|||||
1b 481 GATTAAG 480
UY 541 ACTGAG 600
|||||
1b 541 ACTGAG 600
UY 601 GATTAAG 660
|||||
1b 601 GATTAAG 660
UY 661 GAG 720
|||||
1b 661 GAG 720
UY 721 ATGATGAG 780
|||||
1b 721 ATGATGAG 780
UY 781 AACAAATTCCTGCTTTCTGAGCTCAACCTCTTCTGAACTCAAGCTCTCTGAGAGC 840
|||||
1b 781 AACAAATTCCTGCTTTCTGAGCTCAACCTCTTCTGAACTCAAGCTCTCTGAGAGC 840
UY 841 TCTTCAATGAGAGCTTCAATGAGAGCTGATGCTGCTTCAATGAGAGCTGAGAGC 900
|||||
1b 841 TCTTCAATGAGAGCTTCAATGAGAGCTGATGCTGCTTCAATGAGAGCTGAGAGC 900
UY 901 CTATCTCTCTGCTCAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 960
|||||
1b 901 CTATCTCTCTGCTCAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 960
UY 961 CTGCAAAATCTTGAATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
|||||
1b 961 CTGCAAAATCTTGAATATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
UY 1021 CAAAGAAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
|||||
1b 1021 CAAAGAAATCTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
UY 1081 GTTCAAG 1140
|||||
1b 1081 GTTCAAG 1140
UY 1141 ATTCAGCTCTTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
|||||
1b 1141 ATTCAGCTCTTCTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
UY 1201 GCTCTCTCTGAG 1252
|||||
1b 1201 GCTCTCTCTGAG 1252

RESULT 4
AX323055

LOCUS AX323055 1434 bp DNA linear PAT 07 JAN 2002
DEFINITION Sequence 1 from Patent EP162264.
ACCESSION AX323055
VERSION AX323055.1 GI:18094942
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 Moselman, S. and Dijkema, R.
TITLE Chimeric hormone receptor
JOURNAL Patent: EP 1162264 A 1 12-DEC 2001.
AKZO Nobel N.V. (NL)
FEATURES
source location/Qualifiers
1..1434
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 347 a 373 c 406 g 308 t

ORIGIN

Query Match 99.2%; Score 1247; DB 6; Length 1434;
Best local Similarity 100.0%; Pred. No. 1 6e-271;
Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 ATGAATATACAGATTTCAGCAATGTCACCTGAGAGGTGGGCTGATGGCAGAC 60
DB 1 ATGAATATACAGATTTCAGCAATGTCACCTGAGAGGTGGGCTGATGGCAGAC 60
UY 61 ACAAGGCAAAATGCTGTGGCCAGACCTGGGCGCTTCTGCTTAGTATGTCAGT 120
DB 61 ACAAGGCAAAATGCTGTGGCCAGACCTGGGCGCTTCTGCTTAGTATGTCAGT 120
UY 121 CATTATATACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
DB 121 CATTATATACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
UY 122 CAGTATACACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
DB 122 CAGTATACACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
UY 181 GAACATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 240
DB 181 GAACATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 240
UY 241 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 300
DB 241 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 300
UY 242 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 360
DB 242 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 360
UY 401 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
DB 401 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
UY 402 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
DB 402 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
UY 461 AAAATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 420
DB 461 AAAATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 420
UY 421 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
DB 421 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
UY 422 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
DB 422 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
UY 481 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
DB 481 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
UY 482 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
DB 482 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
UY 541 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 600
DB 541 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 600
UY 601 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
DB 601 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
UY 602 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
DB 602 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
UY 661 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 720
DB 661 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 720

DB 661 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 720

UY 721 ATGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 780
DB 721 ATGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 780
UY 781 AAGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 840
DB 781 AAGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 840
UY 841 TGTGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 900
DB 841 TGTGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 900
UY 901 TGTGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 960
DB 901 TGTGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 960
UY 961 TGTGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1020
DB 961 TGTGATGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1020
UY 1021 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1080
DB 1021 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1080
UY 1081 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1140
DB 1081 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1140
UY 1141 ACCGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1200
DB 1141 ACCGATGCTGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1200
UY 1201 CAGTATACACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1260
DB 1201 CAGTATACACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 1260

RESULT 10
A61586
LOCUS A61586 1460 bp DNA linear PAT 09 MAR 1998
DEFINITION Sequence 4 from Patent W9709348.
ACCESSION A61586
VERSION A61586.1 GI:3715917
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 (bases 1 to 1460)
AUTHORS Kulper, G.G., Emark, F., and Gustafsson, J.
TITLE ORPHAN RECEPTOR
JOURNAL Patent: WO 9709348-A 4 13-MAR-1997;
KAROBIO AB (SE)
FEATURES
source location/Qualifiers
1..1460
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 352 a 479 c 411 g 318 t

ORIGIN

Query Match 99.2%; Score 1247; DB 6; Length 1460;
Best local Similarity 100.0%; Pred. No. 1 6e-271;
Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1 ATGAATATACAGATTTCAGCAATGTCACCTGAGAGGTGGGCTGATGGCAGAC 60
DB 1 ATGAATATACAGATTTCAGCAATGTCACCTGAGAGGTGGGCTGATGGCAGAC 60
UY 61 ACAAGGCAAAATGCTGTGGCCAGACCTGGGCGCTTCTGCTTAGTATGTCAGT 120
DB 61 ACAAGGCAAAATGCTGTGGCCAGACCTGGGCGCTTCTGCTTAGTATGTCAGT 120
UY 121 CATTATATACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
DB 121 CATTATATACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
UY 122 CAGTATACACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
DB 122 CAGTATACACATCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 180
UY 181 GAACATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 240
DB 181 GAACATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 240
UY 241 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 300
DB 241 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 300
UY 242 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 360
DB 242 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 360
UY 401 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
DB 401 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
UY 402 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
DB 402 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 460
UY 461 AAAATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 420
DB 461 AAAATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 420
UY 421 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
DB 421 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
UY 422 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
DB 422 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 480
UY 481 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
DB 481 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
UY 482 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
DB 482 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 540
UY 541 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 600
DB 541 ATGCTGATGCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 600
UY 601 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
DB 601 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
UY 602 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
DB 602 CATAAAATACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 660
UY 661 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 720
DB 661 GATATACCTACCTGATGCGGAGCTTAAAGATGCTGCTGGTGAATAGATGAT 720

661 GAGGCTGAGGCGCCGATGCTGATCAAGCGCCGATGCGCTGATGCGAGCGCTGCT 720
|||||
687 GAGGCTGAGGCGCCGATGCTGATCAAGCGCCGATGCGCTGATGCGAGCGCTGCT 746
|||||
721 ATATGATGATGCTGATGAGCAATGGCGCGAAGAGGCTGGACATGATGAGCTGGCT 780
|||||
747 ATATGATGATGCTGATGAGCAATGGCGCGAAGAGGCTGGACATGATGAGCTGGCT 806
|||||
781 AAGAGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 840
|||||
807 AAGAGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 866
|||||
841 TGTTCATGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 900
|||||
867 TGTTCATGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 926
|||||
901 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 960
|||||
927 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 986
|||||
961 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1020
|||||
987 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1046
|||||
1021 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1080
|||||
1047 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1106
|||||
1081 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1140
|||||
1107 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1166
|||||
1141 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1200
|||||
1167 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1226
|||||
1201 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1247
|||||
1227 GTGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1273
|||||

RESULT 12
AX022458 1460 bp DNA linear PAT 07-SEP-2000
LOCUS AX022458
DEFINITION Sequence 4 from Patent EP0935000.
ACCESSION AX022458
VERSION AX022458.1 GI:10046069
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
FEATURES
Source location/Qualifiers
1..1460
"organism"="Homo sapiens"
"db_xref"="taxon:9606"

Query Match 99.28; Score 1247; DB 6; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1.6e-271;
Matches 1247; conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 60
|||||
DB 27 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 86
|||||
DB 61 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 120
|||||
DB 87 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 146
|||||
DB 121 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 180
|||||

147 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 206
|||||
181 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 240
|||||
207 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 266
|||||
241 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 300
|||||
267 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 326
|||||
301 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 360
|||||
327 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 386
|||||
361 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 420
|||||
387 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 446
|||||
421 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 480
|||||
447 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 506
|||||
481 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 540
|||||
507 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 566
|||||
541 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 600
|||||
567 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 626
|||||
601 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 660
|||||
627 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 686
|||||
661 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 720
|||||
687 GATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 746
|||||
721 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 780
|||||
747 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 806
|||||
781 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 840
|||||
807 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 866
|||||
841 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 900
|||||
867 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 926
|||||
901 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 960
|||||
927 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 986
|||||
961 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1020
|||||
987 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1046
|||||
1021 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1080
|||||
1047 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1106
|||||
1081 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1140
|||||
1107 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1166
|||||
1141 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1200
|||||
1167 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1226
|||||
1201 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1247
|||||
1227 ATGATGCTGAGGCTTGTGAGCTCAAGCTGTTGGAGCAAGTGTGAGCAAT 1273
|||||

PPHVLSRPSAPTEASMMSLDKLAELEWMLTSMWKKLFEVELSLPQRLLESF
 MWEVLMMGLSLLHLLPAPLVLRDQKAVELDLFHLPLATLSREHKL
 QHEEYLVKAMLLINSSMYELVATLQALSSPKLHLNATVLAIVYLASSLSISQ
 QSMRLANLMLLSHVRHSNR-3MHL3MR-7RNVYVTLLELRRAHVDLFEKLT
 GSRSTAEHSKRSKSNFUSQ

BASE COUNT 425 a 463 c 459 g 494 t
 Query Match 99.2% Score 1247; DB 9; Length 1740;
 Best Local Similarity 100.0%; Prod No. 166 271;
 Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCAATTAAAGATATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 60
 108 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 120
 258 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 180
 417 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 240
 61 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 300
 120 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 360
 180 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 420
 240 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 480
 300 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 540
 360 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 600
 420 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 660
 480 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 720
 540 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 780
 600 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 840
 660 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 900
 720 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 960
 780 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1020
 840 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1080
 900 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1140
 960 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1200
 1020 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1260
 1080 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1320
 1140 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1380
 1200 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1440
 1260 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1500
 1320 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1560
 1380 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1620
 1440 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1680
 1500 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1740

108 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1217
 258 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1320
 417 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1423
 61 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1526
 120 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1629
 180 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1732
 240 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1835
 300 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1938
 360 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2041
 420 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2144
 480 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2247
 540 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2350
 600 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2453
 660 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2556
 720 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2659
 780 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2762
 840 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2865
 900 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2968
 960 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3071
 1020 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3174
 1080 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3277
 1140 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3380
 1200 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3483
 1260 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3586
 1320 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3689
 1380 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3792
 1440 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3895
 1500 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3998
 1560 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4101
 1620 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4204
 1680 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4307
 1740 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4410

RESULT 16
 Accession A65127
 Definition Sequence 24 from Patient Ep0798478.
 Accession A65127
 Version A65127.1 GI:4540991
 Keywords unidentifed,
 unidentifed,
 unclassified.
 ORGANISM
 REFERENCES
 1. (bases 1 to 1898)
 Moseleiman, S. and Dijkema, R.
 TITLE
 Estrogen receptor
 JOURNAL
 Patient : Ep 0798478-A 24 01 OCT 1997;
 ARKO NOBEL, NV (NL)
 COMMENT
 of her publication CA 2200423 19970926.
 of her publication AU 1652197 19971002.
 FEATURES
 SOURCE
 1. 1898
 /organism="unidentifed"
 /db_xref="taxon:32644"
 BASE COUNT 452 a 506 c 501 g 459 t
 ORIGIN
 Query Match 99.2% Score 1247; DB 9; Length 1898;
 Best Local Similarity 100.0%; Prod No. 166 271;
 Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATCAATTAAAGATATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 60
 108 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 120
 258 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 180
 417 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 240
 61 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 300
 120 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 360
 180 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 420
 240 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 480
 300 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 540
 360 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 600
 420 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 660
 480 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 720
 540 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 780
 600 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 840
 660 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 900
 720 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 960
 780 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1020
 840 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1080
 900 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1140
 960 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1200
 1020 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1260
 1080 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1320
 1140 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1380
 1200 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1440
 1260 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1500
 1320 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1560
 1380 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1620
 1440 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1680
 1500 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1740

108 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1217
 258 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1320
 417 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1423
 61 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1526
 120 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1629
 180 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1732
 240 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1835
 300 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 1938
 360 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2041
 420 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2144
 480 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2247
 540 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2350
 600 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2453
 660 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2556
 720 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2659
 780 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2762
 840 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2865
 900 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 2968
 960 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3071
 1020 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3174
 1080 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3277
 1140 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3380
 1200 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3483
 1260 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3586
 1320 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3689
 1380 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3792
 1440 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3895
 1500 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 3998
 1560 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4101
 1620 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4204
 1680 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4307
 1740 ATCAATTAAGCAATTCAGAAATGCAATTAATCTTGAAAGCGGGCTGGTGAAGCAAGT 4410


```

|||||
536 GATTAGCATGCGATATCACTATAGGAGTGTGTCTGCTGTAAGCATGTAAAGCTTTT 595
|||||
QY 361 AAAAGACGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 420
|||||
DB 596 AAAAGACGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 555
|||||
QY 421 GATAAAGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 480
|||||
DB 656 GATAAAGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 715
|||||
QY 481 ATGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
|||||
DB 716 ATGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
|||||
QY 541 ATGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
|||||
DB 776 ATGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
|||||
QY 601 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 660
|||||
DB 836 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 895
|||||
QY 661 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 720
|||||
DB 896 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 955
|||||
QY 721 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||||
DB 956 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
|||||
QY 781 AAGAAATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
|||||
DB 1016 AAGAAATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
|||||
QY 841 TGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||||
DB 1076 TGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
|||||
QY 901 TGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
|||||
DB 1136 TGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1195
|||||
QY 961 CTGGAATCTTTCATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
|||||
DB 1196 CTGGAATCTTTCATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1255
|||||
QY 1021 CAGCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
|||||
DB 1256 CAGCAAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1315
|||||
QY 1081 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
|||||
DB 1316 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1375
|||||
QY 1141 AATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
|||||
DB 1376 AATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1435
|||||
QY 1201 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1247
|||||
DB 1436 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1482
|||||

```

```

|||||
Fukaya et al. Melanocyte, Cranial, Vertebral, Intelestom,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Mosseman, S. and Dijkema, R.
TITLE
Chimeric hormone receptor
JOURNAL
Patent: EP 116244-A 24 12-Dec-2001;
Akzo Nobel N.V. (NL)
FEATURES
source
1. 1898
/organism: "Homo sapiens"
/db_xref: "taxon:9606"
BASE COUNT
452 a 506 c 501 g 439 t
ORIGIN
Query: Match 99.2%; Score 1247; DB 6; Length 1898;
Best Local Similarity 100.0%; Pred. No. 1.6e 271;
Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
|||||
DB 246 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 295
|||||
QY 61 ACAAGCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
|||||
DB 296 ACAAGCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355
|||||
QY 121 CAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
|||||
DB 356 CAGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415
|||||
QY 181 GAACACACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
|||||
DB 416 GAACACACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
|||||
QY 241 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
|||||
DB 476 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
|||||
QY 301 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
|||||
DB 536 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
|||||
QY 361 AAAAGACGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 420
|||||
DB 596 AAAAGACGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 655
|||||
QY 421 GATAAAGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 480
|||||
DB 656 GATAAAGATTCAGAGCATATGATTATATTTGCTAGCTAAATAGCTATCAATC 715
|||||
QY 481 ATGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
|||||
DB 716 ATGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
|||||
QY 541 ACTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
|||||
DB 776 ACTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
|||||
QY 601 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
|||||
DB 836 GATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895
|||||
QY 661 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 720
|||||
DB 896 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 955
|||||
QY 721 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
|||||
DB 956 ATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
|||||
QY 781 AAGAAATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
|||||
DB 1016 AAGAAATATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
|||||

```

RESULT 17
 AX123078
 LOCUS AX123078 1898 bp DNA linear PAT 07-JAN-2002
 DEFINITION Sequence 24 from Patent EP1162264.
 ACCESSION AX123078 GI:18093960
 VERSION AX123078.1
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM

FEATURES
Source

Location/Qualifiers
1..1584
/organism="Ovis aries"
/db_xref="taxon:9940"
/issue-type="Ovary"
1..1584
/note="Steroid receptor"
/product="Steroid"
/product="Steroid receptor beta"
/protein_id="A05572.1"
/db_xref="GI:5911384"
/translation="MDKNSPSSLSNPVNCOSLLPPEPPIVLSVSRKPVSV
AVTEYSPAVMVSSTPNSSPDPQMTSPVWPTPCHLSPIALIRGSSILVHPKRS
PWCEPRLPHRYVNHETIKKASNSSTASPSPPSPRIAPPAVSEVAGSRRGV
WSECKAFKRSIUGHNDYICPAINOCITDKNPKSCQAPLPKVEVSVKOSRBR
ERQGYPIVPPQNSPQIHLSTKPAAPMVEKELISLSPPIVITLPAEPH
VMSKSPAFTEASMMSLIKLADKLHMIWAKIIGVELSLVQVLLSKWLE
VLWMLMPSIDHPSKIIPAPVINDPFCVGLIETIMILATISPEELIGER
EYLCVKAILNSMPSKATSOEASUSKRIHLINVAIVLWVIAKSSMSQDSM
PLANIMISHPHNSKCMPEHINMCKYNAVIVYDLEKLNHITLRSNRPVITPSE
RNLAEDSSKSKSPQOQ"

BASE COUNT
497 a 420 c 414 g 353 t

Query Match
Host Local Similarity 87.9%; Score 995; DB 4; Length 1584;
Pred. No. 150-214;
Matches 1085; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

13 ATTCCAGCAATGCTTAATCTGGAAGTGGGCTGCTGGCGAGATACAGAGCCAAAT 72
163 AATTCATATCTCCAAATAACTCAAAAGATGGGCTGGTGCAGAGACAGAGCCAAAT 222
73 GTCTGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 132
223 GATTGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 282
143 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 192
283 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 442
193 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 252
443 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 402
254 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 412
403 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 462
413 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 472
463 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 522
473 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 492
523 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 542
543 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 592
643 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 702
553 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 612
703 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 762
613 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 672
763 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 822
673 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 742

823 CCGTACGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
744 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 792
883 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 942
743 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 852
943 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 902
853 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 912
1003 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1062
913 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 972
1063 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1122
973 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1032
1123 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1182
1033 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1092
1183 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1242
1093 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1152
1243 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1302
1154 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1212
1303 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1362
1213 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1247
1363 GTGATGGGCAACCGCTGGGACCTTCTGCTTATGCTGCTGCTGCTGCTGCTGCTGCT 1397

RESULT 24
AF110402 1611 bp mRNA linear MAM 11-APR-1999
1290US
DEFINITION
Bos taurus breed Holstein estrogen receptor-beta mRNA, complete cds.
ACCESSION
AF110402
VERSION
AF110402.1 GI:4580595
KEYWORDS
Bos taurus.
Bos taurus.
Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1611)
posn:fold, c. S., Yuan, X., Manikkam, M., Calder, M.D., Garverick, H.A.
and Lubahn, D.B.
TITLE
Cloning, sequencing, and localization of bovine estrogen
receptor-beta within the ovarian follicle
JOURNAL
Biol. Reprod. 60 (3), 691-697 (1999)
MEDLINE
99150196
PubMed
10926117
REFERENCE
2 (bases 1 to 1611)
Yuan, X., posn:fold, c. S., Manikkam, M., Calder, M.D., Garverick, H.A.
and Lubahn, D.B.
TITLE
Direct Submission
Submitted (01-DEC-1998) Biochemistry, University of
Missouri-Columbia, 164A Animal Sciences Research Center, 920 E.
Campus Drive, Columbia, MO 65211, USA
Location/Qualifiers
1..1611
/organism="Bos taurus"
/strain="Breed Holstein"
/db_xref="taxon:9919"

WSEED:KAEKESKSTGIBNLYITETAINQULIFENBPKSQAVALERKCYEVMKQVSR
EAGGCTVKKRNSKSHLDLUSKANKNMDHILKVENELSTLSLSPQVATLLEAFPH
VLSRSLPELHSSMMNLSLKAJNKLHMLSWANLHSHLSTLSPQVATLLEAFPH
VLMVHMMSTHPPSKLIPAPVULIDRBRKVSCHLEPHMLATISPEELKSR
EVLVAMILLSSMYPSSAAAPAFSSKRTHTLNAVTHALVWVIAPSGTSQGVSR
LANLMLSHVHJASNKTEHLLNKKKNVIVYDILLEMNAHLIKNSLVDKSR
SRESESEKESKQVQAO*

BASE COUNT 996 a 424 c 411 g 450 t
ORIGIN

Query Match 76.6% Score 96.4 DB 4: Length 1581

Host Label Similarity 86.9% Prod. No. 2.7e-207
Matches 1072 Conserved 92 Mismatches 160 Indels 4 Gaps 1

13 ATTCCAGCAATCTTGTGAAAGCTGGAGCTGGAGCAAGCAAGCCCAAT 72
14 ATTTAGCAATCTTGTGAAAGCTGGAGCTGGAGCAAGCAAGCCCAAT 72
15 AATTAGCAATCTTGTGAAAGCTGGAGCTGGAGCAAGCAAGCCCAAT 72
16 AATTAGCAATCTTGTGAAAGCTGGAGCTGGAGCAAGCAAGCCCAAT 72
17 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
18 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
19 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
20 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
21 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
22 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
23 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
24 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
25 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
26 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
27 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
28 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
29 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
30 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
31 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
32 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
33 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
34 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
35 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
36 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
37 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
38 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
39 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
40 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
41 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
42 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
43 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
44 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
45 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
46 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
47 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
48 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
49 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
50 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
51 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
52 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
53 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
54 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
55 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
56 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
57 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
58 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
59 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
60 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
61 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
62 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
63 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
64 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
65 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
66 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
67 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
68 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
69 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
70 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
71 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
72 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
73 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
74 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
75 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
76 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
77 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
78 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
79 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
80 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
81 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
82 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
83 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
84 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
85 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
86 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
87 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
88 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
89 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
90 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
91 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
92 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
93 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
94 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
95 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
96 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
97 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
98 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
99 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
100 GTCTGTGAAAGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142

913 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1064 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
974 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1124 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1034 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1184 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1094 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1240 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1154 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1400 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1213 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142
1460 GTGATGCTGGAGCTGGAGCTGGAGCAAGCAAGCCCAAT 142

RESULT 27
AF267736 1581 bp mRNA linear MAR 15 MAR 2002
DEFINITION
AF267736 AF267736.1 G1:3184420
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa.
Sus scrofa.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
REFERENCE
1 (bases 1 to 1581)
Lavoie,H.A., Deslaurier,D.C., Gilling-McLain,C. and Hui,Y.Y.
Cloning and characterization of porcine ovarian estrogen receptor
beta isoforms
Mol. Reprod. 66 (4), 616-624 (2002)
MEDLINE
21858415
11870066
PUBMED
2 (bases 1 to 1581)
Lavoie,H.A.
REFERENCES
TITLE
Direct Submission
JOURNAL
of South Carolina School of Medicine, Bldg 1, Columbia, SC 29208,
USA

FEATURES

Source

1..1581
/organism "Sus scrofa"
/db_xref "taxon:9843"
/feature "type "genetic"
1..1581
/gene "Er beta"
1..1581
/gene "Er beta"
/gene "Er beta"
/product "estrogen receptor beta"
/protein_id "AAK15151.1"

Notes

Translation: "MILKNSKSTGIBNLYITETAINQULIFENBPKSQAVALERKCYEVMKQVSR
EAGGCTVKKRNSKSHLDLUSKANKNMDHILKVENELSTLSLSPQVATLLEAFPH
VLSRSLPELHSSMMNLSLKAJNKLHMLSWANLHSHLSTLSPQVATLLEAFPH
VLMVHMMSTHPPSKLIPAPVULIDRBRKVSCHLEPHMLATISPEELKSR
EVLVAMILLSSMYPSSAAAPAFSSKRTHTLNAVTHALVWVIAPSGTSQGVSR
LANLMLSHVHJASNKTEHLLNKKKNVIVYDILLEMNAHLIKNSLVDKSR
SRESESEKESKQVQAO"

Query Match:		74.8%	Score 940.2	DB 10	Length 1450
Best Local Similarity:		84.9%	Pred. No. 8/6/2002		
Matches 1053:		Conservative 0	Mismatches 186	Indels 0	Gaps 0
UY	1	ATGATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	60		
DB	217	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	276		
UY	61	ACAAGGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	120		
DB	277	ACAAGGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	346		
UY	121	CAGTATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	180		
DB	347	CAGTATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	396		
UY	141	GAACATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	240		
DB	397	GAACATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	456		
UY	241	GCTGATTCGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	300		
DB	457	GCTGATTCGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	516		
UY	301	GATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	460		
DB	517	GATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	576		
UY	361	AAAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	420		
DB	577	AAAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	636		
UY	421	GATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	480		
DB	637	GATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	696		
UY	481	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	540		
DB	647	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	706		
UY	541	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	600		
DB	757	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	816		
UY	601	CGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	660		
DB	817	CGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	876		
UY	661	GATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	720		
DB	877	GATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	936		
UY	721	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	780		
DB	937	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	996		
UY	741	AAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	800		
DB	947	AAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1006		
UY	841	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	900		
DB	1057	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1116		
UY	901	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	960		
DB	1117	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1176		
UY	961	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1020		
DB	1177	ATGATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1236		
UY	1021	GAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1080		

DB	1237	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1296		
UY	1081	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1140		
DB	1297	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1356		
UY	1141	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1200		
DB	1357	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1416		
UY	1201	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1260		
DB	1417	CAAGGACATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1476		

FEATURE	32	END57439	END57439	2555 bp	GENE	LOC 12-100-1996
DEFINITION		Rattus norvegicus estrogen receptor beta mRNA, complete cds.				
ACCESSION		U57439				
VERSION		U57439.1				
KEYWORDS						
SOURCE		Rattus norvegicus.				
ORGANISM		Rattus norvegicus.				

REFERENCE	1 (bases 9 to 2555)	Kulper, G.G., Eumark, E., Pello-Huikko, M., Nilsson, S., and Gustafsson, J.A.
TITLE	Cloning of a novel receptor expressed in rat prostate and ovary	
GENERAL	66234066	
MEDLINE	8650195	
REFERENCE	2 (bases 1 to 2555)	Eumark, E.L.
TITLE	Direct Submission	
GENERAL	Submitted (03-MAY-1996)	
REFERENCE	3 (bases 1 to 2555)	Eumark, E.L.
TITLE	Submitted (03-MAY-1996)	
GENERAL	Submitted (03-MAY-1996)	
REFERENCE	4 (bases 1 to 2555)	Eumark, E.L.
TITLE	Submitted (03-MAY-1996)	
GENERAL	Submitted (03-MAY-1996)	

FEATURES	1..2555	location/Qualifiers
source	1..2555	organism="Rattus norvegicus"
		strain="Sprague Dawley"
		sex="male"
		tissue="prostate"
		note="nuclear receptor superfamily member"
		product="estrogen receptor beta"
		protein_id="AAC52602.1"
		db_xref="db:1373281"
		translation="MTFTSPAVMNTSVGTSNIDGGVRLSTSPNVLFTSCHLSPL
		ATFTVSLIYAPKSPWLEASLEHLLPVPLILKPISSSTASVSLSPNVRAB
		FEPCVSDVAGSYHYGVMSCEGKAFKPKSLGGHNDYICPAINQTIIRKPKSGARF
		LRKKEVEMVWKGSPKPGGYVIVPQPSSEVYHCLSKARKNGHAPRKELLSL
		SPEOLVITLFAEPNVILVSRSPFTFASMMISUTKLAKELVHMGKRIIGVEY
		LSLIDVRLISCHMVEVIMVGMIMMSIDHPKLIIPAPVILIDREKCVGSLLEFLM
		ILATTSRRELKIDREKPYCYKAMILINSWPIASANGPAPSPSKTHLNAVTL
		VWVAFSLSSQSGSVIRANILMLISHPHLSNGQHLISMKKNVIVYDILLEML
		NAHTLQKSSISSESTERSKRSKNSONLSQ"
		note="c" encodes DNA binding domain"
		note="g" encodes DNA binding domain"

BASE COUNT	630 a	652 c	660 g	614 t
ORIGIN				
Query Match	74.8%	Score 947	DB 10	Length 2555
Best Local Similarity	84.9%	Pred. No. 2/10-201		
Matches 1051	Conservative 0	Mismatches 190	Indels 0	Gaps 0

UY	1	ATGATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	60		
DB	1117	ATGATATACGATTCACGACATGTCATATCTTAAAGTGGAGTGGAGGACACAC	1176		

GenCore version 5.1.6
Copyright (c) 1994 - 2003 Computer Ltd.

OM nucleotide - nucleotide search, using SW model

Run on: June 15, 2003 23:19:19 : Search time 298.206 seconds

(without alignments)
9822 000 Million cell updates/sec

File: US-08-826-361A-20

Perfect score: 1257

Sequence: 1 ATGATTCAGCATTCCTCAT

Scoring table: IDENTITY, MISC

Gapop 10.0, Gapext 1.0

Search: 218530 seqs, 11059949 residues

Total number of hits satisfying chosen parameters: 4370470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 120 summaries

```
Database: N_Genoseq_101002.*
1: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1960.DAT.*
2: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1961.DAT.*
3: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1962.DAT.*
4: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1963.DAT.*
5: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1964.DAT.*
6: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1965.DAT.*
7: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1966.DAT.*
8: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1967.DAT.*
9: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1968.DAT.*
10: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1969.DAT.*
11: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1970.DAT.*
12: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1971.DAT.*
13: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1972.DAT.*
14: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1973.DAT.*
15: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1974.DAT.*
16: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1975.DAT.*
17: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1976.DAT.*
18: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1977.DAT.*
19: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1978.DAT.*
20: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1979.DAT.*
21: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA1980.DAT.*
22: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA2001A.DAT.*
23: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA2001B.DAT.*
24: /SITS2/qpdata/qgenseq/qgenseq-comb1/NA2002.DAT.*
```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length DB	ID	Description
1	1257	100.0	1257	18	Human oestrogen re
2	1257	100.0	1257	24	Human oestrogen re
3	1249	99.4	1251	18	Human oestrogen re
4	1249	99.4	1251	24	Human oestrogen re
5	1247	99.2	1434	18	Human oestrogen re
6	1247	99.2	1434	24	Human oestrogen re
7	1247	99.2	1460	18	Human oestrogen re
8	1247	99.2	1467	20	Human oestrogen re
9	1247	99.2	1686	20	Human oestrogen re

10	1247	99.2	1740	22	AAF59897	Human oestrogen re
11	1247	99.2	1898	24	AAI72144	Human oestrogen re
12	1247	99.2	1336	18	AAI88115	Human oestrogen re
13	1247	99.2	2011	24	AAI88142	Human oestrogen re
14	1247	99.2	2011	24	AAI88142	Human oestrogen re
15	1247	99.2	2433	24	AAI88125	Human oestrogen re
16	1247	99.2	1460	20	AAI88125	Human oestrogen re
17	1247	99.2	1460	20	AAI88125	Human oestrogen re
18	1247	99.2	1460	20	AAI88125	Human oestrogen re
19	1247	99.2	1460	20	AAI88125	Human oestrogen re
20	1247	99.2	1460	20	AAI88125	Human oestrogen re
21	1247	99.2	1460	20	AAI88125	Human oestrogen re
22	1247	99.2	1460	20	AAI88125	Human oestrogen re
23	1247	99.2	1460	20	AAI88125	Human oestrogen re
24	1247	99.2	1460	20	AAI88125	Human oestrogen re
25	1247	99.2	1460	20	AAI88125	Human oestrogen re
26	1247	99.2	1460	20	AAI88125	Human oestrogen re
27	1247	99.2	1460	20	AAI88125	Human oestrogen re
28	1247	99.2	1460	20	AAI88125	Human oestrogen re
29	1247	99.2	1460	20	AAI88125	Human oestrogen re
30	1247	99.2	1460	20	AAI88125	Human oestrogen re
31	1247	99.2	1460	20	AAI88125	Human oestrogen re
32	1247	99.2	1460	20	AAI88125	Human oestrogen re
33	1247	99.2	1460	20	AAI88125	Human oestrogen re
34	1247	99.2	1460	20	AAI88125	Human oestrogen re
35	1247	99.2	1460	20	AAI88125	Human oestrogen re
36	1247	99.2	1460	20	AAI88125	Human oestrogen re
37	1247	99.2	1460	20	AAI88125	Human oestrogen re
38	1247	99.2	1460	20	AAI88125	Human oestrogen re
39	1247	99.2	1460	20	AAI88125	Human oestrogen re
40	1247	99.2	1460	20	AAI88125	Human oestrogen re
41	1247	99.2	1460	20	AAI88125	Human oestrogen re
42	1247	99.2	1460	20	AAI88125	Human oestrogen re
43	1247	99.2	1460	20	AAI88125	Human oestrogen re
44	1247	99.2	1460	20	AAI88125	Human oestrogen re
45	1247	99.2	1460	20	AAI88125	Human oestrogen re
46	1247	99.2	1460	20	AAI88125	Human oestrogen re
47	1247	99.2	1460	20	AAI88125	Human oestrogen re
48	1247	99.2	1460	20	AAI88125	Human oestrogen re
49	1247	99.2	1460	20	AAI88125	Human oestrogen re
50	1247	99.2	1460	20	AAI88125	Human oestrogen re
51	1247	99.2	1460	20	AAI88125	Human oestrogen re
52	1247	99.2	1460	20	AAI88125	Human oestrogen re
53	1247	99.2	1460	20	AAI88125	Human oestrogen re
54	1247	99.2	1460	20	AAI88125	Human oestrogen re
55	1247	99.2	1460	20	AAI88125	Human oestrogen re
56	1247	99.2	1460	20	AAI88125	Human oestrogen re
57	1247	99.2	1460	20	AAI88125	Human oestrogen re
58	1247	99.2	1460	20	AAI88125	Human oestrogen re
59	1247	99.2	1460	20	AAI88125	Human oestrogen re
60	1247	99.2	1460	20	AAI88125	Human oestrogen re
61	1247	99.2	1460	20	AAI88125	Human oestrogen re
62	1247	99.2	1460	20	AAI88125	Human oestrogen re
63	1247	99.2	1460	20	AAI88125	Human oestrogen re
64	1247	99.2	1460	20	AAI88125	Human oestrogen re
65	1247	99.2	1460	20	AAI88125	Human oestrogen re
66	1247	99.2	1460	20	AAI88125	Human oestrogen re
67	1247	99.2	1460	20	AAI88125	Human oestrogen re
68	1247	99.2	1460	20	AAI88125	Human oestrogen re
69	1247	99.2	1460	20	AAI88125	Human oestrogen re
70	1247	99.2	1460	20	AAI88125	Human oestrogen re
71	1247	99.2	1460	20	AAI88125	Human oestrogen re
72	1247	99.2	1460	20	AAI88125	Human oestrogen re
73	1247	99.2	1460	20	AAI88125	Human oestrogen re
74	1247	99.2	1460	20	AAI88125	Human oestrogen re
75	1247	99.2	1460	20	AAI88125	Human oestrogen re
76	1247	99.2	1460	20	AAI88125	Human oestrogen re
77	1247	99.2	1460	20	AAI88125	Human oestrogen re
78	1247	99.2	1460	20	AAI88125	Human oestrogen re
79	1247	99.2	1460	20	AAI88125	Human oestrogen re
80	1247	99.2	1460	20	AAI88125	Human oestrogen re
81	1247	99.2	1460	20	AAI88125	Human oestrogen re
82	1247	99.2	1460	20	AAI88125	Human oestrogen re

[illegible]

```

RES011.2
AA172146
ID AA172146 standard; cDNA; 1257 BP.
XX
XX
AC AA172146;
XX
XX
BI 25-MAR-2002 (first entry)
XX
XX
DE cDNA encoding ER splice variant ER-alpha.
XX
XX
KW DNA binding domain; DBD; ligand binding domain; LBD; chimeric receptor;
KW estrogen receptor; ER; chromosome 14; ER-alpha; ER-beta; exon 8; PCR;
KW estradiol; nuclear receptor; ER-positive; ER-negative; amplified; protein;
KW polymerase chain reaction; Af-2 region; splice variant; ss.
XX
XX
OS Homo sapiens.
XX
XX
Key location/Qualifiers
FH l.1257
FI CDS
FI /*tag=
FI /product="ER-alpha"
XX
XX
EP162264-A2.
XX
XX
PN 12-DEC-2001.
XX
XX
PE 25-MAR-1997; 2001FP-0206021.
XX
XX
PE 26-MAR-1996; 96EP-0200820.
XX
XX
PE 22-NOV-1994; 96EP-0203284.
XX
XX
PE 25-MAR-1997; 97EP-0200904.
XX
XX
FA (ALKU ) AKZO NOBEL NV.
XX
XX
PI MosseIman S., D[]kema R.
XX
XX
WP1: 2002-08414/12.
XX
XX
DR P-PSDB; AAB47836.
XX
XX

```

PT New isolated chimera receptor comprising a DNA binding domain and/or
 PE DNA binding domain of a new estrogen receptor, for identifying
 PT functional ligands or hormonal analogs for the receptor
 XX
 PS Example B; Page 23; 45pp; English.
 XX
 CC The sequences given in AAI72146 and AAI72153 encode splice variants
 CC of a novel estrogen receptor (ER). The gene encoding this new ER is
 CC located on chromosome 14 and has a different tissue distribution from
 CC classical ER. This ER also has two orphan ER's, ER-alpha and ER-beta.
 CC These orphan receptors have estrogen receptor related structure but do
 CC not appear to be able to bind estradiol or other ER ligands. These
 CC splice variants differ from the parent ER around exon 8. One clone
 CC contains exon 8B through alternative splicing. This causes an immediate
 CC termination of the reading thereby creating a truncation at the carboxy
 CC terminus. The other splice variant contains an alternative exon 8, exon
 CC 8C, which encodes two C-terminal amino acids and then has a stop codon.
 CC The splice variant proteins do not contain the A6-2 region and therefore
 CC probably lack the ability to modulate transcription of target genes in
 CC a ligand dependant manner. However they may be able to interfere with
 CC the functioning of the WT classical ER or the novel ER of the invention,
 CC either by heterodimerisation or by occupying estrogen response elements
 CC or by interactions with other transcription factors.
 CC
 XX
 XX Sequence 1257 BP; 297 A; 327 C; 355 G; 278 T; 0 other;
 SO
 Query Match 100.0%; Score 1257; DB 24; Length 1257;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1257; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CY 1 ATGAATTCACGATTCCTCCGCAATGTCTACTACTTGGAAAGTGGGCTGTGGTGGACGC 60
 DB 1 ATGAATTTAAATGATTCGCGCAATGTCTACTACTTGGAAAGTGGGCTGTGGTGGACGC 60
 CY 61 ACAGAGCCCAAAAGTGTGTGTGGCAACAGCGGGGACCTTTCTTACTGGTGCATAC 120
 DB 61 ACAGAGCCCAAAAGTGTGTGTGGCAACAGCGGGGACCTTTCTTACTGGTGCATAC 120
 DB 61 ACAGAGCCCAAAAGTGTGTGTGGCAACAGCGGGGACCTTTCTTACTGGTGCATAC 120
 CY 121 CAGTATATACATCTGTATGTGGAAAGCTCAAAAGATCTGTGTGTAAATGAATGTGTA 180
 DB 121 CAGTATATACATCTGTATGTGGAAAGCTCAAAAGATCTGTGTGTAAATGAATGTGTA 180
 CY 121 CAGTATATACATCTGTATGTGGAAAGCTCAAAAGATCTGTGTGTAAATGAATGTGTA 180
 DB 121 CAGTATATACATCTGTATGTGGAAAGCTCAAAAGATCTGTGTGTAAATGAATGTGTA 180
 CY 181 GAACCAACGCTACAGCTGTAAACAGACACACATGTGAAGAGAGGATATAGAGAGCTGTG 240
 DB 181 GAACCAACGCTACAGCTGTAAACAGACACACATGTGAAGAGAGGATATAGAGAGCTGTG 240
 CY 181 GAACCAACGCTACAGCTGTAAACAGACACACATGTGAAGAGAGGATATAGAGAGCTGTG 240
 DB 181 GAACCAACGCTACAGCTGTAAACAGACACACATGTGAAGAGAGGATATAGAGAGCTGTG 240
 CY 241 GGTATATCTGTATATCTGTAAATGTAAATGTAAATGTAAATGTAAATGTAAATGTAAAT 300
 DB 241 GGTATATCTGTATATCTGTAAATGTAAATGTAAATGTAAATGTAAATGTAAATGTAAAT 300
 CY 401 GATTATGTATGTGATATCTACTATGTGAGATCTGTGTGTAAAGATGTAAAGCTTTT 360
 DB 401 GATTATGTATGTGATATCTACTATGTGAGATCTGTGTGTAAAGATGTAAAGCTTTT 360
 CY 461 AAAGCAAGCATTCGAAGCATATATATATTGTGTCACTATCAAAATAGAGTACATC 420
 DB 461 AAAGCAAGCATTCGAAGCATATATATATTGTGTCACTATCAAAATAGAGTACATC 420
 CY 421 GATTAATAAAGCTGT 480
 DB 421 GATTAATAAAGCTGT 480
 CY 481 ATGGTGAAGT 540
 DB 481 ATGGTGAAGT 540
 CY 541 ATGTGTGAAGT 600
 DB 541 ATGTGTGAAGT 600
 CY 601 CGAGT 660
 DB 601 CGAGT 660


```

UY 721 ATGATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 780
DB 721 ATGATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 780
UY 781 AAGAAAGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 840
DB 781 AAGAAAGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 840
UY 841 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 900
DB 841 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 900
UY 901 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 960
DB 901 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 960
UY 961 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1020
DB 961 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1020
UY 1021 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1080
DB 1021 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1080
UY 1081 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1140
DB 1081 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1140
UY 1141 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1200
DB 1141 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1200
UY 1201 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1260
DB 1201 TGTGATGATGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 1260

```

RESULT 4
ID AA172153 standard: cDNA: 1251 bp.
XX
AC AA172153:
XX
UT 25-MAR-2002 (first entry)
XX
DE cDNA encoding ER splice variant ER-beta.
XX
KW DNA binding domain; DBD; ligand binding domain; ER; chimeric receptor;
KW estrogen receptor; ER; chromosome 14; ER-alpha; ER-beta; exon 8; ER;
KW estradiol; nuclear receptor; progesterone receptor; amplification;
KW polymerase chain reaction; AF-2 region; splice variant; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 1..1251
FT /tag = a
FT /product = "ER-beta"
XX
FN EP162264-A2.
XX
PD 12-DEC-2001.
XX
PE 25-MAR-1997; 2001EP 0202021.
XX
FR 26-MAR-1996; 96EP-0200820.
FR 22-NOV-1996; 96EP-0201284.
FR 25-MAR-1997; 97EP-0200903.
XX
FA (AKU) AKZO NOBEL NV.
XX
PI Moselman S. Dijkema R.

```

DB WP1: 2002-084414/12.
DB P PSDB: AAH47837.
XX
PT New isolated chimeric receptor comprising a DNA binding domain and/or
PT ligand binding domain of a new estrogen receptor, for identifying
PT functional ligands or hormonal analogs for the receptor
XX
PS Example B: Page 13: 35pp: English.
XX
CC The sequences given in AA172146 and AA172153 encode splice variants
CC of a novel estrogen receptor (ER). The gene encoding this new ER is
CC located on chromosome 14 and has a different tissue distribution from
CC classical ER. This ER also has two orphan ERs, ER-alpha and ER-beta.
CC These orphan receptors have estrogen receptor related structure but do
CC not appear to be able to bind estradiol or other ER ligands. These
CC splice variants differ from the parent ER around exon 8. One clone
CC contains exon 8B through alternative splicing. This causes an immediate
CC termination of the reading thereby creating a truncation at the carboxy
CC terminus. The other splice variant contains an alternative exon 8, exon
CC 8C, which encodes two C-terminal amino acids and then has a stop codon.
CC The splice variant proteins do not contain the AF-2 region and therefore
CC probably lack the ability to modulate transcription of target genes in
CC a ligand dependant manner. However they may be able to interfere with
CC the functioning of the WT classical ER or the novel ER of the invention,
CC either by heterodimerisation or by occupying estrogen response elements
CC or by interactions with other transcription factors.
XX
S0 Sequence 1251 bp: 297 A: 324 C: 354 G: 276 T: 0 other:

```

Query Match 99.48; Score 1249; DB 24; Length 1251;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

UY 1 ATGATTAATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 60
DB 1 ATGATTAATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 60
UY 61 ATAAAGGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 120
DB 61 ACAGAGGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 120
UY 121 CAGTATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 180
DB 121 CAGTATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 180
UY 181 GAGCAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 240
DB 181 GAGCAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 240
UY 241 GAGCAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 300
DB 241 GAGCAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 300
UY 301 GATTAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 360
DB 301 GATTAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 360
UY 361 AAAAGAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 420
DB 361 AAAAGAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 420
UY 421 GATTAAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 480
DB 421 GATTAAGCTGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 480
UY 481 ATGATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 540
DB 481 ATGATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 540
UY 541 ATGATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 600
DB 541 ATGATGATGCTGCTGCAAAATGTCGCGCAAGAGAGTGTGTACATCATATGATGCTGCGGCTG 600

```


[illegible]

functional ligands or hormonal analogs for the receptor

Example A: Page 12: 35pp; English.

XX
 XX This sequence encodes a novel estrogen receptor (ER). The gene encoding
 CC this new ER is located on chromosome 14 and has a different tissue
 CC distribution from classical ER. This ER also has two orphan ER's,
 CC ER-alpha and ER-beta. These orphan receptors have estrogen receptor
 CC related structure but do not appear to able to bind estradiol or other
 CC ER ligands. The DNA binding domain (DBD) and ligand binding domain (LBD)
 CC from this ER may be used in the chimeric receptor of the invention
 CC which also has an N-terminal domain. The chimeric receptor, or DNA
 CC encoding it, is useful in a screening assay for identification of new
 CC drugs. Similar chimeric receptors comprising the LBD of the new ER,
 CC and also comprising the ERD and an N-terminal domain derived from
 CC another nuclear receptor e.g., progesterone receptor, are useful for the
 CC screening of compounds to identify new ligands or hormone analogs which
 CC are able to activate the new ER. Chimeric receptors comprising a DBD
 CC of the new ER, and LBD and an N-terminal domain derived from
 CC another nuclear receptor, can be used to identify new ligands or
 CC hormone analogs for the nuclear receptors.

XX
 XX Sequence 1434 BP; 347 A; 373 C; 406 G; 308 T; 0 other:

GCFFY Match: 99.2%; Score 1247; ES 24; Length 1414;
 Host: Local Similarity 100.0%; Prod. No. 0;
 Matches 1247; Conservative 0; Mismatches 0; Indels 0; Gaps 0

07 1 ATGATATTACGACTTCCGACGACATCTGCTACCTGGAAAGCGTGGCTGGCTGGCGGAAAGC 60
 08 1 ATCAATATTACGACTTCCGACGACATCTGCTACCTGGAAAGCGTGGCTGGCGGAAAGC 60
 09 61 ACACAGTCCAAATCTCTTGTGTGTCGAAACATCTGGAGACTTCTCTCTTAACTGTTATATCC 120
 10 61 ACACAGTCCAAATCTCTTGTGTGTCGAAACATCTGGAGACTTCTCTCTTAACTGTTATATCC 120
 11 61 ACACAGTCCAAATCTCTTGTGTGTCGAAACATCTGGAGACTTCTCTCTTAACTGTTATATCC 120
 12 121 CAGTTATACATCTGTATATCGGAACTTCAAAAAGCTTCCCTGGTGTGAACTCAATATGCTTA 180
 13 121 CAGTTATACATCTGTATATCGGAACTTCAAAAAGCTTCCCTGGTGTGAACTCAATATGCTTA 180
 14 121 CAGTTATACATCTGTATATCGGAACTTCAAAAAGCTTCCCTGGTGTGAACTCAATATGCTTA 180
 15 181 GAACATACCTTACTCTGTAAACATGAAGACATCTGAAGAGAGAGTATAGTGGAAACCTTTC 240
 16 181 GAACATACCTTACTCTGTAAACATGAAGACATCTGAAGAGAGAGTATAGTGGAAACCTTTC 240
 17 181 GAACATACCTTACTCTGTAAACATGAAGACATCTGAAGAGAGAGTATAGTGGAAACCTTTC 240
 18 241 GCGAGGCTCTGTATCTGATCGAGTTCGAAAGAGAGATGCTCATCTGTGCTGCTTCAGAC 300
 19 301 GATTATCCATCGCATATATCATATGAGATCTGGTTCGTGTGAAGCATGTAAGAGCTTTT 360
 20 301 GATTATCCATCGCATATATCATATGAGATCTGGTTCGTGTGAAGCATGTAAGAGCTTTT 360
 21 361 AAAAGAGATTTAAATGAATATATGATGATTATTTGTCGATCTTAATAATCATGTTAAAT 420
 22 361 AAAAGAGATTTAAATGAATATATGATGATTATTTGTCGATCTTAATAATCATGTTAAAT 420
 23 421 GATTAAAAATGGCGGTGAAGATTCGCAGAGTCTGTGCACTTCGCAAGCTGTACGAATGGCA 480
 24 421 GATTAAAAATGGCGGTGAAGATTCGCAGAGTCTGTGCACTTCGCAAGCTGTACGAATGGCA 480
 25 421 GATTAAAAATGGCGGTGAAGATTCGCAGAGTCTGTGCACTTCGCAAGCTGTACGAATGGCA 480
 26 481 ATGGTGAAGATGGAGTGTGGAAGAGAGATGAGGATGCGGCTGCTGGTGGTGAAGATGAAGA 540
 27 481 ATGGTGAAGATGGAGTGTGGAAGAGAGATGAGGATGCGGCTGCTGGTGGTGAAGATGAAGA 540
 28 541 AGTGGTGAAGATGGAGTGTGGAAGAGAGATGAGGATGCGGCTGCTGGTGGTGAAGATGAAGA 600
 29 541 AGTGGTGAAGATGGAGTGTGGAAGAGAGATGAGGATGCGGCTGCTGGTGGTGAAGATGAAGA 600
 30 601 GGAATGAGAGATGCTGTCTGTGAAGAGAGTGAAGTGGAGAGCTAGAGCTTACCTGCTG 660
 31 601 GGAATGAGAGATGCTGTCTGTGAAGAGAGTGAAGTGGAGAGCTAGAGCTTACCTGCTG 660
 32 661 GGAATGAGAGATGCTGTCTGTGAAGAGAGTGAAGTGGAGAGCTAGAGCTTACCTGCTG 660
 33 661 GGAATGAGAGATGCTGTCTGTGAAGAGAGTGAAGTGGAGAGCTAGAGCTTACCTGCTG 660
 34 661 GGAATGAGAGATGCTGTCTGTGAAGAGAGTGAAGTGGAGAGCTAGAGCTTACCTGCTG 660

[illegible]

protein

claim 26: Page 68-73: 75pp: Japanese.

The invention relates to a method for examining the interaction of a test substance with a hormone receptor protein. The method involves contacting a test substance with a hormone receptor protein (e.g., an oestrogen receptor) which is labelled with an optical marker (e.g., a fluorescent protein), where the marker is capable of undergoing a change in its optical properties when the receptor binds a ligand. The optical signal generated in the presence of the test compound is then compared with that generated in the absence of the test compound to determine whether the test compound is interacting with the hormone receptor protein. The invention also encompasses the labelled hormone receptor protein, the gene encoding it, and vectors and host cells comprising such genes. The method of the invention is used for the investigation and diagnosis of hormonal disorders, especially those associated with sex hormones e.g., suppression of ovulation. The present sequence represents cDNA encoding human oestrogen receptor beta (ER beta).

Sequence 1740 BP: 425 A: 463 C: 459 G: 493 T: 0 other:

Query Match 99.2% Score 1247: DB 22: Length 1740:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 1247: Conservative 0: Mismatches 0: Indels 0: Gaps 0

1 ATGATATACACATTATCCAGTCAGTACATGCTACTCTGGAGAGATGGAGCTGGTGGAGAT 60
|||||
258 ATGATATACACATTATCCAGTAAATCTCAACTTTCAGATTTGAGCTGATGGAGAGAT 117
61 ACAGAGCCGAATGTGTGTGGTCAACACTGGAGACCTTCTCTCTTACTGGTCAATCAT 120
|||||
318 ACAGAGCCGAATGTGTGTGGTCAACACTGGAGACCTTCTCTCTTACTGGTCAATCAT 177
121 CAGTTTTCACATCTGTATGCTGCAACCTTCAAAAAGATCTCTGGTGTGAGACACATGCTA 180
|||||
178 CAGTTTTCACATCTGTATGCTGCAACCTTCAAAAAGATCTCTGGTGTGAGACACATGCTA 237
181 GAACATACCTTACTCTTAATATGAGATAGATCTGAAGAGAGAGATTAATGAGATCTGAT 240
|||||
438 GAACATACCTTACTCTTAATATGAGATAGATCTGAAGAGAGAGATTAATGAGATCTGAT 497
241 GCGAGGCTGTCTTACCTGCTCAGAGTTCAAGAGAGAGATCTGCTGCTGCTGTGATCAT 300
|||||
498 GCGAGGCTGTCTTACCTGCTCAGAGTTCAAGAGAGAGATCTGCTGCTGCTGTGATCAT 357
401 GATTACGCAATCCGAGATATCATATATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
|||||
558 GATTACGCAATCCGAGATATCATATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
361 AAAAGAGAGATTCAAGGATATATATATATATTTCTCTCATATTAATATAGTGTATATAT 420
|||||
618 AAAAGAGAGATTCAAGGAGATATATATATATTTCTCTCATATTAATATAGTGTATATAT 477
421 GATTAAGAAATCGCGTAGAGAGCTGCGCAGAGCTGGCGCACTCTGGAAGTCTTACGAAAGTGGGA 480
|||||
678 GATTAAGAAATCGCGTAGAGAGCTGCGCAGAGCTGGCGCACTCTGGAAGTCTTACGAAAGTGGGA 737
481 ATGCTCAAGATGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
|||||
738 ATGCTCAAGATGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 797
541 AGTGTGCAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
|||||
798 AGTGTGCAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
601 GCAATGCGCAGAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
|||||
858 GCAATGCGCAGAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
661 GAGAGTGAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
|||||
918 GAGAGTGAATGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977

QY	66	1248	QY	721	Db	QY	901	Db	QY	961	Db	QY	1021	Db	QY	1081	Db	QY	1141	Db	QY	1201	Db	QY	1261	Db	QY	1321	Db	QY	1381	Db	QY	1441	Db	QY	1501	Db	QY	1561	Db	QY	1621	Db	QY	1681	Db	QY	1741	Db	QY	1801	Db	QY	1861	Db	QY	1921	Db	QY	1981	Db	QY	2041	Db	QY	2101	Db	QY	2161	Db	QY	2221	Db	QY	2281	Db	QY	2341	Db	QY	2401	Db	QY	2461	Db	QY	2521	Db	QY	2581	Db	QY	2641	Db	QY	2701	Db	QY	2761	Db	QY	2821	Db	QY	2881	Db	QY	2941	Db	QY	3001	Db	QY	3061	Db	QY	3121	Db	QY	3181	Db	QY	3241	Db	QY	3301	Db	QY	3361	Db	QY	3421	Db	QY	3481	Db	QY	3541	Db	QY	3601	Db	QY	3661	Db	QY	3721	Db	QY	3781	Db	QY	3841	Db	QY	3901	Db	QY	3961	Db	QY	4021	Db	QY	4081	Db	QY	4141	Db	QY	4201	Db	QY	4261	Db	QY	4321	Db	QY	4381	Db	QY	4441	Db	QY	4501	Db	QY	4561	Db	QY	4621	Db	QY	4681	Db	QY	4741	Db	QY	4801	Db	QY	4861	Db	QY	4921	Db	QY	4981	Db	QY	5041	Db	QY	5101	Db	QY	5161	Db	QY	5221	Db	QY	5281	Db	QY	5341	Db	QY	5401	Db	QY	5461	Db	QY	5521	Db	QY	5581	Db	QY	5641	Db	QY	5701	Db	QY	5761	Db	QY	5821	Db	QY	5881	Db	QY	5941	Db	QY	6001	Db	QY	6061	Db	QY	6121	Db	QY	6181	Db	QY	6241	Db	QY	6301	Db	QY	6361	Db	QY	6421	Db	QY	6481	Db	QY	6541	Db	QY	6601	Db	QY	6661	Db	QY	6721	Db	QY	6781	Db	QY	6841	Db	QY	6901	Db	QY	6961	Db	QY	7021	Db	QY	7081	Db	QY	7141	Db	QY	7201	Db	QY	7261	Db	QY	7321	Db	QY	7381	Db	QY	7441	Db	QY	7501	Db	QY	7561	Db	QY	7621	Db	QY	7681	Db	QY	7741	Db	QY	7801	Db	QY	7861	Db	QY	7921	Db	QY	7981	Db	QY	8041	Db	QY	8101	Db	QY	8161	Db	QY	8221	Db	QY	8281	Db	QY	8341	Db	QY	8401	Db	QY	8461	Db	QY	8521	Db	QY	8581	Db	QY	8641	Db	QY	8701	Db	QY	8761	Db	QY	8821	Db	QY	8881	Db	QY	8941	Db	QY	9001	Db	QY	9061	Db	QY	9121	Db	QY	9181	Db	QY	9241	Db	QY	9301	Db	QY	9361	Db	QY	9421	Db	QY	9481	Db	QY	9541	Db	QY	9601	Db	QY	9661	Db	QY	9721	Db	QY	9781	Db	QY	9841	Db	QY	9901	Db	QY	9961	Db	QY	10021	Db	QY	10081	Db	QY	10141	Db	QY	10201	Db	QY	10261	Db	QY	10321	Db	QY	10381	Db	QY	10441	Db	QY	10501	Db	QY	10561	Db	QY	10621	Db	QY	10681	Db	QY	10741	Db	QY	10801	Db	QY	10861	Db	QY	10921	Db	QY	10981	Db	QY	11041	Db	QY	11101	Db	QY	11161	Db	QY	11221	Db	QY	11281	Db	QY	11341	Db	QY	11401	Db	QY	11461	Db	QY	11521	Db	QY	11581	Db	QY	11641	Db	QY	11701	Db	QY	11761	Db	QY
----	----	------	----	-----	----	----	-----	----	----	-----	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----	-------	----	----

xx oestrogen receptor gene and protein polymorphisms useful for diagnosis
 pt of individuals at risk of developing bone disorders -
 xx
 xx Disclosure: Figure 37 245ppt. English.
 ps

The invention relates to a novel isolated peptide comprising or consisting of an amino acid sequence selected from an amino acid sequence of a variant estrogen receptor protein (e.g., EPReta), or a fragment of 10 amino acids), antibodies against them, nucleic acids encoding them (including vectors for transforming cells). The gene for human EPReta is located on chromosome 6q, 25, 1. The variants are encoded by stable nucleotide polymorphisms (SNP). The variant peptides and proteins can be used in assays to determine the biological activity of the protein, to raise antibodies, as a reagent in assays designed to quantitatively determine levels of the protein in biological fluids, to identify compounds that modulate receptor activity and to screen compounds for the ability to stimulate or inhibit interaction between the receptor protein and a target molecule that normally interacts with the receptor protein e.g., oestrogen. The antibody can be used to isolate the protein, to assess expression in disease states e.g. cardiovascular disease and autoimmune disease (e.g., systemic lupus erythematosus, arthritis, rheumatism and osteoarthritis), osteoporosis, breast cancer and endometrial cancer. In addition the antibodies can be used in pharmacogenomic analysis and inhibiting protein function, e.g. blocking the binding of the oestrogen receptor protein to a binding partner such as a ligand. The nucleic acids encoding the proteins can be used as probes, primers, chemical intermediates and in biological assays. The present sequence encodes human EPReta.

50 Sequence 2011 BP; 470 A; 541 C; 561 G; 439 T; 0 other;

Query Match	99.28	Score 1247	DB 22	Length 1011
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1247	Conservative 0	Mismatches 0	Indels 0	Gaps 0

[illegible]


```

721 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 780
745 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 804
761 AATGAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 840
805 AATGAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 864
841 TGTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 900
865 TGTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 924
901 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 960
925 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 984
961 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1020
985 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1044
1021 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1080
1045 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1104
1081 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1140
1105 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1164
1141 AGGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1200
1165 AGGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1224
1201 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1241
1225 GTGATGCTGCTGATGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 1265

```

RESULT 21

AAK4482

AAK4482

25 JUN 1999 (first entry)

Murine mdr beta 3 clone DNA sequence.

Estrogen receptor beta; estrogen receptor beta; ER beta; ds.

Mus sp.

Key

Location/Qualifiers

1..1600

/*tag a

/product "complete estrogen receptor beta"

/note "no stop codon given"

W0900171 A1.

04 FEB 1999.

28 JUL 1998: 9880 081540.

00 JUL 1997: 9708 0014210.

28 JUL 1997: 9708 0054869.

(OYVA) UNIT VALUE.

Barton R., Chandon P., Barton R., Bopard S., Garolier J.

W01: 1999 14285412.

P F50H: AAY04433.

```

XX Model murine estrogen receptor beta genes and splice variants
PT used for treating estrogen receptor beta mediated disorders
XX
XX Example 1: Fig 1b-1: 89pp; English.
XX
XX The present sequence encodes the murine complete estrogen receptor beta
XX (mER beta) also designated mdr beta 3). The protein can be used to
XX screen for agents that modulate or block the interaction between the
XX ER beta protein and physiologically, stilbene or triphenylethylene
XX (anti)estrogens. The antibody which binds the protein can be used in
XX assays to detect activation of the ER beta protein and can be used for
XX subcellular localization of the protein. The complementary nucleic acid
XX can be used to inhibit expression of the ER beta c nucleic acid.
XX
XX Sequence 1704 bp; 424 A; 457 C; 448 G; 475 T; 0 other;

```

Query Match 68.98; Score 866.6; 106 20; Length 1704;

Post Local Similarity 80.8%; Prod. No. 4.2e 218;

Matches 1047; Conservative 0; Mismatches 194; Indels 54; Gaps 1;

```

1 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 60
1111 1111 11111111 1 111 111111111111 11 1111
217 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 276
1111 1111 11111111 1 111 111111111111 11 1111
61 AATGAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 120
1111 1111 11111111 1 111 1111 11111111 11 11
277 GAAAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 336
1111 1111 11111111 1 111 1111 11111111 11 11
121 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 180
11 111 111 1111 111111111111 111111111111 11
327 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 396
11 111 111 1111 111111111111 111111111111 11
181 GAAAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 240
1111 1111 111111111111 1111 1111 1111 1111 1111
397 GAAAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 456
1111 1111 111111111111 1111 1111 1111 1111 1111
241 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 300
1111 1111 1111 1111 111111111111 11111111 111111
457 GAAAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 516
1111 1111 1111 1111 111111111111 11111111 111111
401 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 460
1111 1111 1111 1111 1111 111111111111 11111111 111111
517 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 576
1111 1111 1111 1111 1111 111111111111 11111111 111111
461 AATGAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 420
1111 1111 111111111111 1111 1111 1111 1111 1111
577 AATGAGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 480
1111 1111 111111111111 1111 1111 1111 1111 1111
421 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 480
11 11 111111 11 1111111111111111 1111111111 111
637 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 696
11 11 111111 11 1111111111111111 1111111111 111
481 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 540
1111 1111 1111 1111 1111 1111111111 11 11 11111111
697 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 756
1111 1111 1111 1111 1111 1111111111 11 11 11111111
541 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 600
1111 1111 1111 1111 1111 1111111111 11 11 11 1111
757 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 816
1111 1111 1111 1111 1111 1111111111 11 11 11 1111
601 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 660
11 111 1111 111111 11 1111 1111111111 1111111111 111
817 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 876
11 111 1111 1111 1111 1111111111 1111111111 111111
661 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 720
11 111111 11 111111 11 1111 1111 1111111111111111
877 GATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 936
1111 1111 1111 1111 1111 11111111 1111 11111111 11111111
721 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 780
1111 1111 1111 1111 1111 11111111 1111 11111111 11111111
937 ATGATGATGCTGCTGATGAGTGGGAGAAAGCAATGCTGAAATGATGATGAGGAGG 996
1111 1111 1111 1111 1111 11111111 1111 11111111 11111111

```

```

25 781 AAGAACATTCGCGGCTTTGGAGCTGAGCTGTCAGCAAGTGGGCTCTTGGAGAG 840
    ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 997 AAGAAATTCGCTGGCTTTGGAGCTGAGCTGTCAGCAAGTGGGCTCTTGGAGAG 1056
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 841 TCTTATGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1057 TCTTATGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 901 CTGATCTTCTGTCAGATCTTCTGTCAGATCTTCTGTCAGATCTTCTGTCAGAT 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1117 CTGATCTTCTGTCAGATCTTCTGTCAGATCTTCTGTCAGATCTTCTGTCAGAT 1176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 981 -----AGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1177 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 967 ATCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 926
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1237 ATCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 1027 GATATCTCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1297 GATATCTCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 1087 GCGACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1357 GCGACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1416
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 1147 GCTTGGCTTGGCTGATTCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1417 GCTTGGCTTGGCTGATTCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 1207 GCTTGGCTTGGCTGATTCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1267
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1477 GCTTGGCTTGGCTGATTCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 22

AA24935
ID AAX24935 standard; cDNA: 1704 bp

AC AAX24935;

21-JUN-1999 (first entry)

Mouse complete oestrogen receptor beta clone mEPbeta-3

Oestrogen receptor: ER beta-3; ER beta-3; mouse; tumour; diagnosis;

therapy: ds

Mus musculus.

Key: local/lin/qualifiers

FT exon 1149..1203

FT /tag= a

FT /number= 5R

W0905170-AT.

04-FEB-1999.

28-JUL-1998: 98WD-US15539.

30-JUL-1997: 97US-0054210

28-JUL-1997: 97US-0054869.

(UYVA) UNIV VALE.

Denton R, Fraslon C, Galien R, Rische-Rigon M;

WP1: 1999-142852/12.

DR P-BSDH: AAW98128.

XX

Human oestrogen receptor beta-3 - useful in treating and monitoring
tumours

Example 1: Fig 1b: 89pp; English.

The present sequence is the coding region of the murine complete
oestrogen receptor beta-3 (ep beta-3) cDNA clone mEP beta-3, which
encodes a 567 amino acid protein (see AAW98128). The gene contains
9 exons, including the newly described exon 5b, which is not found
in ER beta-1. Exon 5b is located in the ligand binding domain of
mEP beta-3, and probably plays a significant role in mEP beta-3
function. The mEP beta-3 clone was twice isolated using 2
separate procedures: RT-PCR of mRNA, and amplification from a
mouse embryonic stem cell genomic DNA library. 2 Alternatively
spliced forms of mEP beta-3 were also identified: mEP beta-1 (see
AAW98128), which lacks exon 5b; and mEP beta-2 (see AAW98128), which
lacks exon 3. A third isoform, ER beta-4 (see AAW98128) was
identified in rat ovaries. The human exon 5b region (see AAX24934)
has also been identified. Based on these sequences, the invention
provides methods for identifying agents that block or augment
ER beta-3 mediated transcriptional regulation, methods to determine
whether ER beta-3 or its isoforms is being expressed in tissues or
cells, and methods of identifying and using agents that block the
transcriptional regulation of genes by ER beta-3 or its isoforms,
which in turn modulate other biological and pathological
processes. Gene therapy expression of ER beta-3 is envisaged.

Sequence 1704 BP; 424 A; 458 C; 447 G; 375 T; 0 other;

Query Match: 68.8%; Score 865; DB 20; Length 1704;

Best local Similarity: 80.8%; Pred. No. 8.4e-218;

Matches 1046; Conservative 0; Mismatches 195; Indels 54; Gaps 1;

```

25 1 ATGATATTAAGATTTCCAGCAATGTCACCTACTACTGTAAGAGTGGGCTTGTCGACAC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 ATGAACTACAGCTGTTCCAGCAAGTACCGATACCTGAGGAGGCTGGCTGTTGGCA 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 61 ACAAGCCCAATATGTTTGGCAATACCTGGGACGCTTCTGTTAGTGGTACAC 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 277 GCAAGCCCAATATGTTTGGCAATACCTGGGACGCTTCTGTTAGTGGTACAC 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 121 CAGTTATCAGATCTGATGCGAAGCTCAAAAGAGCTGGTGGTGAAGCAATGCTGA 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 CAGTTATCAGATCTGATGCGAAGCTCAAAAGAGCTGGTGGTGAAGCAATGCTGA 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 181 GAGCAGACCTTATGTTTAAATAGCAATCTGAAAGAGAGGTTATGCTGAACCTGG 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 397 GAGCAGACCTTATGTTTAAATAGCAATCTGAAAGAGAGGTTATGCTGAACCTGG 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 241 GCGAGCCGTTATGCTGGTCAAGTTCGCAAGAGGAGGATGCTGCTGCTGCAAC 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 457 GCGAGCCGTTATGCTGGTCAAGTTCGCAAGAGGAGGATGCTGCTGCTGCAAC 316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 301 GATTACGATACGATATGATATGAGTCTGGTGTGTAAGAGAGTGAAGCTTTT 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 517 GATTACGATACGATATGATATGAGTCTGGTGTGTAAGAGAGTGAAGCTTTT 376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 361 AAAGAGAGATTCGCAAGATATGATATGCTGCAATGCTGCAATGCTGCAATG 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 577 AAAGAGAGATTCGCAAGATATGATATGCTGCAATGCTGCAATGCTGCAATG 436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 421 GATTAAGAGAGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 637 GATTAAGAGAGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 696
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 481 ATGGTGAAGTGGCTGCGAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 697 ATGGTGAAGTGGCTGCGAGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 756
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25 541 AGTGGCAAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 757 AGTGGCAAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


XX AAD1046:
 AC 18-JUN-2002 (first entry)
 DE Medaka fish oestrogen receptor cDNA.
 XX Medaka fish: oestrogen receptor; transgenic; thrombosis; therapeutic
 XX endocrine disrupting chemical; bioassay system; ss.
 XX Oryzias latipes.
 XX Key location/qualifiers
 FT CDS 211..1948
 FT /feature a
 FT /product: "Medaka fish oestrogen receptor protein"
 XX
 XX EP1180684-A1.
 XX 20-FEB-2002.
 XX 27-JUN-2001: 2001EP-0115472.
 XX 17-AUG-2000: 2000JP-0247729.
 XX (OYH-) UNIV HIROSHIMA.
 XX Yamashita I.
 XX WP1: 2992 229967/29.
 XX P-PSDB: ABE19430.
 XX
 XX New polynucleotide from Oryzias latipes, useful for producing
 PT transgenic medaka fish having sensitivity to a very small amount of
 PT estrogen, which is useful for testing estrogen-like compounds
 PT disrupting chemicals.
 XX
 XX Claim 1: Page 9-16; 23pp; English.
 XX
 XX The present sequence is medaka fish (Oryzias latipes) oestrogen receptor
 CC cDNA. The polynucleotide is useful for producing transgenic medaka fish
 CC having sensitivity to a very small amount of estrogen or a medaka fish
 CC having one or more thereof. The highly estrogen-sensitive medaka fish
 CC is useful for testing estrogen-like endocrine disrupting chemicals,
 CC for elucidating developmental mechanism of thrombosis, or as a bioassay
 CC system for developing a therapeutic agent for thrombosis.
 XX
 XX Sequence 2764 BP; 649 A; 746 C; 725 G; 654 T; 0 other.

Query Match 27.0%; Score 338.8; PB 24; Length 2764;
 Host Local Similarity 61.8%; Pred. No. 5,46-79;
 Matches 632; Conservative 0; Mismatches 342; Indels 48; Gaps 4;

QY 274 GATCTGCACTTCTGCGCTCTGCGGATTAAGGATGCGGATATGCTATGAGATCTGG 433
 DB 619 GATCTGCACTTCTGCGCTCTGCGGATTAAGGATGCGGATATGCTATGAGATCTGG 678
 QY 334 TCTGTGAGCATGTAAAGCTTTTAAAGAGACATTAAGACATTAAGATTAAT 493
 DB 679 TCTGTGAGCATGTAAAGCTTTTAAAGAGACATTAAGACATTAAGATTAAT 749
 QY 494 TGTGAGCATGTAAAGCTTTTAAAGAGACATTAAGACATTAAGATTAAT 453
 DB 739 TGTGAGCATGTAAAGCTTTTAAAGAGACATTAAGACATTAAGATTAAT 799
 QY 454 GCACTGCACTTCTGCGCTCTGCGGATTAAGGATGCGGATATGCTATGAGATCTGG 494
 DB 799 GCACTGCACTTCTGCGCTCTGCGGATTAAGGATGCGGATATGCTATGAGATCTGG 858
 QY 495 -----CTGCGGATTAAGGATGCGGATATGCTATGAGATCTGGATTAAGATTAAG 549
 DB 859 GCACTTCTGCGCTCTGCGGATTAAGGATGCGGATATGCTATGAGATCTGGATTAAG 918

QY 560 GAGTACTGCT-----ATTGCTGCGGATTAAGGATTAAGGATTAAGGATTAAG 594
 DB 919 GGTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 978
 QY 595 GCGGCGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 654
 DB 979 GGTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1038
 QY 655 CTCTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 708
 DB 1039 CTCTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1098
 QY 709 ACTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 768
 DB 1099 ACTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1158
 QY 769 ATGATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 828
 DB 1159 ATGATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1218
 QY 829 CTCTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 888
 DB 1219 CTCTGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1278
 QY 889 GATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 948
 DB 1279 GATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1338
 QY 949 GTAGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1008
 DB 1339 GTAGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1398
 QY 1009 TTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1068
 DB 1399 TTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1458
 QY 1069 ATGATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1125
 DB 1459 ATGATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1518
 QY 1125 TTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1185
 DB 1519 ATGATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1578
 QY 1186 GATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1245
 DB 1579 GATGAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAG 1638
 QY 1246 AG 1247
 DB 1639 AG 1640
 RESULT 25
 ID A168126 standard; cDNA: 1824 bp.
 AC A168126;
 DE Lepomis centrarchus oestrogen receptor encoding cDNA. SEQ ID No 5.
 DE Lepomis centrarchus oestrogen receptor; ss.
 XX Lepomis centrarchus.
 XX Key location/qualifiers
 FT CDS 74..1822
 FT /feature a
 FT /product: "oestrogen receptor"
 XX
 XX JP2001197890-A.

UY	27	AAAGATATGTCATCTTGGCTGGCTGGCTGAGGATATGAGTAAAGATGATGATGAT	430
DB	1		60
UY	601	AAGTAAATGAGCGTTTGTGCTGTGTGTGTAAGATATATGCTTGTGCTATGACATG	660
DB	1		60
UY	431	TACTGTCTGAGAGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	490
DB	1		60
DB	661	TACTGTCTGAGAGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	720
UY	491	ATTCTGTGATATTAAGAGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGAT	450
DB	1		60
DB	721	ATTCTGTGATATTAAGAGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGAT	780
UY	451	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	510
DB	1		60
DB	781	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	840
UY	511	TGTGGGATAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	570
DB	1		60
DB	841	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	900
UY	571	AAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	630
DB	1		60
DB	901	AAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	960
UY	630	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	690
DB	1		60
DB	961	AAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1020
UY	690	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	750
DB	1		60
DB	1021	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1080
UY	750	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	810
DB	1		60
DB	1081	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	1140
UY	810	ATTCTGTGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATATAT	870
DB	1		60
DB	1141	ATTCTGTGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATATAT	1200
UY	870	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	930
DB	1		60
DB	1201	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1260
UY	930	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	990
DB	1		60
DB	1261	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
UY	990	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	1050
DB	1		60
DB	1321	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	1380
UY	1050	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1110
DB	1		60
DB	1381	ATTCTGTGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATATAT	1440
UY	1110	ATTCTGTGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATATAT	1170
DB	1		60
DB	1441	ATTCTGTGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATATAT	1500
UY	1170	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1230
DB	1		60
DB	1501	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1560
UY	1230	TATGAGATTCGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATAT	1290
DB	1		60
DB	1561	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1620
UY	1290	GAGCGCTTTGTGGGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1350
DB	1		60
DB	1621	ATTCTGTGATATTAAGAGCGCTTTTATAAAGAGAGATTAAGAGATATATATAT	1680

```

RESULT 27
AAA92174
ID AAA92174 standard; cDNA; 1728 BP.
XX
XX
XX AAA92174:
XX
XX 05-JAN-2001 (first entry)
XX
XX Oryzias latipes: oestrogen receptor encoding cDNA SP010 N012.
XX
XX Oryzias latipes: oestrogen receptor; ds.
XX
XX Oryzias latipes.
XX
XX JF2000201688-A.
XX
XX 25-JUL-2000.
XX
XX 06-APR-1999; 94JP-0098787.
XX
XX 10-NOV-1998; 98JP-0319465.
XX
XX (SUMO) (SUMITOMO CHEM CO LTD.
XX
XX WP1: 2000-567950/53.
XX
XX P-PSDB; AAB20897.
XX
XX An estrogen receptor gene and its application.
XX
XX Claim 3; Page 11-13; 23pp; Japanese.
XX
XX The present sequence encodes an estrogen receptor derived from
XX Oryzias latipes. Also described are: (1) a vector comprising the
XX estrogen receptor gene; (2) a transformant prepared by introducing
XX the estrogen receptor gene or vector from (1) into a host cell;
XX (3) a method for the preparation of an estrogen receptor comprising
XX culturing the transformant from (2) to produce the estrogen receptor;
XX and (4) a method for the evaluation of estrogen receptor activating
XX ability of a chemical substance in which the chemical substance is
XX reacted with a transformant prepared by introducing a reporter gene
XX connected downstream of a transcription controlling region containing
XX an estrogen response sequence and the above estrogen receptor gene to
XX an estrogen nondependent host cell. The transformant can be used for
XX the evaluation of estrogen receptor-activating ability of a chemical
XX substance.
XX
XX Sequence 1728 BP; 378 A; 514 C; 497 G; 439 T; 0 other:
XX
Query Match 26.4%; Score 152.4; Dh 21; Length 1728;
Host Local Similarity 61.4%; Prod. No. 2.2e-77;
Matches 628; Conservative 0; Mismatches 346; Indels 48; Gaps 4.
XX
XX 274 GATGCTACTCTTGGAGCTGCTGAGCATTAAGCATATGAGATATGATATGATGATG 333
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX 409 GATACGCGTTTGTGGGCTGTGTGAGGAGCAGCAGCAGCTGTGGAGTACGATGAGTGG 468
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX
XX 334 TTTGTGCAAGCATGTAAAGCCCTTTTAAAGAAACATTTAAAGCAATATATATAT 393
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX 469 TCTTGGAGAGGCTGTAAAGGCGCTCTTCTTAAAGAGAACTTAAAGCTTAAATGATATAG 528
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX 494 TGTGCTACATCAAAATGACATGTCAATATCAATAAAGAGAGAGTAAATGATGAGAGCTGT 453
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX 529 TATCCATCCATCAATGATGCTCAATATTCATCAAGAACTCAAGGCAAGGCTTGTAGGCTTGG 588
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX
XX 454 CAACTTGGAAATGTTTAAATTTGGAATTTGAAATTTGATGCTTCTCTCTCTCTCTCTCTCT 494
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX 589 CGCTTAAAGAAATGTTTAAATTTGGAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAA 648
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
XX 649 CCGATTTTAAAGGCTGCTCAAGAGGAGAGGCTGTGTGATGATGATGATGATGATGATGATG 708
XX 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

```


XX 11-APR-2002.
 PD
 XX 29-SEP-2001: 2001W-1B02246.
 PF
 XX 03-OCT-2000: 2000ER-0012570.
 PR
 XX 11-MAY-2001: 2001US-0853034.
 PA (ADEFER) ADEFERIN ASS-6C LEV RECH EN GENETIC.
 XX Chabon P, Met-Layer D;
 XX WPI: 2002-463217/49.
 DR P-PSDB: ABR6478.
 XX
 PT A transgenic mouse, useful in screening for medicaments for the
 PT treatment of e.g., diabetes or skin cancers, comprises a fusion protein
 PT between a recombinase Cre, and a modified ligand binding domain of the
 PT nuclear estrogen receptor alpha
 XX
 PS Disclosure: Page 128-130: 14pp: English.

CC The present sequence is the coding sequence of the human nuclear
 CC estrogen receptor alpha (1). The invention relates to a non-human
 CC mammalian organism, especially a transgenic mouse, characterized in
 CC that at least one cell comprises: (1) a fusion protein formed from
 CC a recombinase (Cre), a hinge region (preferably derived from human
 CC (1)), and a modified ligand binding domain (LBD) of a nuclear
 CC estrogen receptor, such as (1). Its treatment or variant; and (2)
 CC one or more genes or DNA sequences of interest belonging to the
 CC genome of the organism, into which one or more recognition sites
 CC of the recombinase protein are inserted. The (1) LBD domain is
 CC preferably mediated by a G521R, G400Y, or M438L/G544A mutation.
 CC The fusion protein has negligible, or even zero, recombinase
 CC activity in the presence of a natural ligand such as estradiol,
 CC but recombinase activity is induced by a small quantity of a
 CC synthetic ligand that has no estrogenic activity, e.g. 4-methyl-
 CC or 4-hydroxylamoxifen. The mammalian organism or its cells, such
 CC as epidermal cells, hepatocytes or adipocytes, are useful in
 CC carrying out a spatiotemporally controlled site-specific
 CC recombination of a DNA sequence of interest in its natural
 CC chromatin environment. It is also used in screening of medicaments
 CC for pathological conditions associated with an alteration of the
 CC expression and/or function of the DNA sequence of interest, such as
 CC skin cancer, inflammation, diabetes, diabetes, obesity, or in
 CC promoting hepatic regeneration.

XX Sequence 1788 BP; 418 A; 515 C; 508 G; 447 T; 0 other;

XX Query Match: 25.6%; Score 322; BB 24; Length 1788;
 XX Host Local Similarity 60.7%; Pred. No. 1,28-74;
 XX Matches 617; Conservative 0; Mismatches 360; Indels 59; Gaps 4;

UY 271 AAGGATGCTGACCTTCTGGGCTGCTGACGATGATGCGATGATGACGATG 330
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 538 AAGGATGCTGACCTTCTGGGCTGCTGACGATGATGCGATGATGACGATG 597
 UY 331 TGGTCTGTGAAGCATGATGATGATGATGATGATGATGATGATGATGAT 390
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 598 TGGTCTGTGAAGCATGATGATGATGATGATGATGATGATGATGATGAT 657
 UY 491 ATTGTCAGCTACCAATGATGATGATGATGATGATGATGATGATGATGAT 450
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 658 ATTGTCAGCTACCAATGATGATGATGATGATGATGATGATGATGATGAT 717
 UY 451 TGGTCTGTGAAGCATGATGATGATGATGATGATGATGATGATGATGAT 510
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 718 TGGTCTGTGAAGCATGATGATGATGATGATGATGATGATGATGATGAT 777
 UY 511 TGTGTG-----GTACGCTCTTGTGGGAGAGAGAGAGAGAGAGAGAGAG 558
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 778 AATGTCAGCTACCAATGATGATGATGATGATGATGATGATGATGATGAT 837

UY 559 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 838 GTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 UY 608 -----GGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 898 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
 UY 658 CTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 958 TGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
 UY 715 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1018 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
 UY 775 TGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1078 TGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
 UY 835 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1138 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
 UY 895 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1198 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
 UY 955 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1258 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317
 UY 1015 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1318 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377
 UY 1075 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1378 GAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
 UY 1132 AAGGATGCTGACCTTCTGGGCTGCTGACGATGATGCGATGATGACGATG 1191
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1438 AAGGATGCTGACCTTCTGGGCTGCTGACGATGATGCGATGATGACGATG 1497
 UY 1192 AAGGATGCTGACCTTCTGGGCTGCTGACGATGATGCGATGATGACGATG 1247
 1b 111 111 111 111 111 111 111 111 111 111 111 111 111
 1498 AAGGATGCTGACCTTCTGGGCTGCTGACGATGATGCGATGATGACGATG 1554
 DB
 RESULT 32
 AAT06873
 ID AAT06873 standard; DNA; 4963 BP.
 XX
 AC AAT06873;
 XX
 DT 13-MAR-1996 (first entry)
 XX
 DE PRCMV*-1 promoter and human estrogen receptor.
 KW PRCMV*-1 promoter and human estrogen receptor.
 KW Telomeric responsive promoter; PRCMV*-1; human cytomegalovirus;
 KW Transcription activation; cytokine; cytokine; cytokine;
 KW ds: ss.
 OS Chimeric synthetic;
 OS Chimeric homo sapiens.
 XX
 PN US5464758-A.
 XX
 PD 07-NOV-1995.
 XX
 PY 14-JUN-1995; 9408-007626.

775 TGAGGTAAGAAATTCGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 844
 1543 TGGGTAAGAAATTCGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1602
 835 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 894
 1603 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1662
 895 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 954
 1663 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1722
 955 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1014
 1723 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1782
 1015 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1074
 1783 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1842
 1075 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1131
 1843 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1902
 1132 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1194
 1903 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1962
 1192 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1247
 1963 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 2018

RESULT 36
 ID AX01370 standard; DNA: 4963 BP.
 AX01370:
 19-APR-1999 (first entry)
 PHMV-1 rabbit proestosterone fusion sequence.
 tTA transactivator; transgenic mouse; tet operator-linked gene;
 tetracycline-controlled transactivator fusion protein;
 transcriptional regulatory element; ds.
 Synthetic.
 US859410-A.
 12 JAN 1999.
 07-JUN-1995: 950S-0481970.
 07-JUN-1995: 950S-0481970.
 14-JUN-1995: 940S-0076427.
 14-JUN-1994: 940S-0260452.
 (BAD) BASE AG.
 Bulard H, Gosson M, Salfeld JG, Voss JM;
 WPL: 1999-152108/13.
 Transgenic mouse - with tetracycline-dependent gene suppression
 Discovered: Fig 10; 78pp; English.
 This sequence represents a fusion gene between the PHMV-1 DNA
 sequence (containing the tet operator sequence and the proMV-1
 promoter) and the rabbit proestosterone receptor coding sequence. The

invention relates to a transgenic mouse having a transgene integrated
 into the genome of the mouse and also having a tet operator-linked gene
 in the genome of the mouse, where the transgene comprises a
 transcriptional regulatory element functional in cells of the mouse
 operatively linked to a polynucleotide sequence encoding a
 tetracycline-controlled transactivator fusion protein (tTA). The fusion
 protein comprises a tet repressor operably linked to a polypeptide which
 directly or indirectly activates transcription of the tet operator-linked
 gene in eukaryotic cells; the tet operator-linked gene controls a
 detectable and functional phenotype on the mouse when expressed in cells
 of the mouse; the transgene is expressed in cells of the mouse at a level
 sufficient to produce amounts of the fusion protein that are sufficient
 to activate transcription of the tet operator-linked gene; and in the
 absence of tetracycline or a tetracycline analogue in the mouse, the
 fusion protein binds to the tet operator-linked gene and activates
 transcription of the tet operator-linked gene such that the tet
 operator-linked gene is expressed at a level sufficient to control the
 detectable and functional phenotype on the mouse, where the level of
 expression of tet operator-linked gene can be down-modulated by
 administration of tetracycline or a tetracycline analogue to the mouse. The
 mice are useful as conditional knockout animals for studying human
 diseases.

Sequence 4963 BP, 1259 A, 1902 C, 1278 G, 1126 T, 0 other.
 Query Match 25.63; Score 322; DB 20; Length 4963;
 Best Local Similarity 60.78; Pred. NO. 186-74;
 Matches 617; Conservative 0; Mismatches 260; Indels 49; Gaps 4;

271 AGGAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 330
 1963 AGGAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1062
 431 TGGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 490
 1063 TGGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1122
 891 ATTGCTGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 950
 1123 ATTGCTGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1182
 451 TGGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 510
 1183 TGGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1242
 511 TGGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 568
 1243 AGGAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1302
 559 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 617
 1303 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1362
 608 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 667
 1363 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1422
 658 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 714
 1423 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1482
 715 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 774
 1483 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1542
 775 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 834
 1543 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1602
 835 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 894
 1603 GAGAGCTTTGAGAGCTTTTGAGAGCTATGCTTTGAGTAAATGCGGCTCTTGG 1662

Genome version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

us nucleic - nucleic search, using sw model

Run on: June 16, 2003, 00:20:22 : Search time 60,100 seconds
(without alignments)
6512 699 Million cells (31467200)

Title: US-08-826-361a-20

Sequence: 1 ATGATTAAGATTCGAC GATATGAAATTCGATTA 1257

Perfect score: 1257

Scoring table: IDENTITY_NMC GAPOP 10.0 : GAPEXT 1.0

Search: 441862 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 822724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 120 summaries

Database : Issued_Patents_NA :
1: /seq2_6/prodata/1/seq2_6/seq1
2: /seq2_6/prodata/1/seq2_6/seq2
3: /seq2_6/prodata/1/seq2_6/seq3
4: /seq2_6/prodata/1/seq2_6/seq4
5: /seq2_6/prodata/1/seq2_6/seq5
6: /seq2_6/prodata/1/seq2_6/seq6

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	1247	99.2	US-08-836-620A-4	Sequence 4, Appl
2	1247	99.2	US-09-139-617-2	Sequence 2, Appl
3	1247	99.2	US-09-561-741A-2	Sequence 2, Appl
4	937	74.5	US-08-836-620A-1	Sequence 1, Appl
5	922.5	73.4	US-08-836-620A-6	Sequence 1, Appl
6	322	25.6	US-08-836-620A-1	Sequence 1, Appl
7	322	25.6	US-08-836-620A-6	Sequence 1, Appl
8	322	25.6	US-08-836-620A-1	Sequence 1, Appl
9	322	25.6	US-08-836-620A-6	Sequence 1, Appl
10	322	25.6	US-08-836-620A-1	Sequence 1, Appl
11	320.4	25.5	US-09-041-886-134	Sequence 1, Appl
12	320.4	25.5	US-09-041-886-134	Sequence 1, Appl
13	216.8	17.2	US-08-836-620A-1	Sequence 1, Appl
14	216.8	17.2	US-08-836-620A-6	Sequence 1, Appl
15	216.8	17.2	US-08-836-620A-1	Sequence 1, Appl
16	113.8	9.1	US-09-040-508-1	Sequence 1, Appl
17	113.8	9.1	US-09-040-508-1	Sequence 1, Appl
18	95.2	7.6	US-09-041-886-15	Sequence 1, Appl
19	89	7.1	US-08-836-620A-1	Sequence 1, Appl
20	89	7.1	US-08-836-620A-6	Sequence 1, Appl
21	89	7.1	US-08-836-620A-1	Sequence 1, Appl
22	89	7.1	US-08-836-620A-6	Sequence 1, Appl
23	89	7.1	US-08-836-620A-1	Sequence 1, Appl
24	84.8	6.7	US-08-479-913E-1	Sequence 1, Appl
25	84.8	6.7	US-08-479-913E-1	Sequence 1, Appl
26	81	6.4	US-08-629-939-1	Sequence 1, Appl
27	81	6.4	US-08-629-939-1	Sequence 1, Appl

28	80.8	6.4	1576	6	US-04-42-1	Patent No. 526042-1
29	78	6.2	2989	6	US-02-66-1	Patent No. 522606
30	74.4	5.9	1797	4	US-09-277-078-1	Sequence 1, Appl
31	74.4	5.9	2658	4	US-08-592-383-3	Sequence 1, Appl
32	74.4	5.9	2928	4	US-08-095-728B-4	Sequence 4, Appl
33	74.4	5.9	2928	5	PCT-US92-0228A-4	Sequence 4, Appl
34	74.4	5.9	2940	2	US-08-592-383-1	Sequence 1, Appl
35	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
36	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
37	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
38	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
39	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
40	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
41	74.4	5.9	2940	6	US-08-592-383-1	Sequence 1, Appl
42	72.8	5.8	2067	4	US-09-106-194-11	Sequence 1, Appl
43	71.2	5.7	387	1	US-08-087-151-1	Sequence 1, Appl
44	71.2	5.7	387	1	US-08-087-151-1	Sequence 1, Appl
45	71.2	5.7	1662	1	US-08-336-408B-5	Sequence 1, Appl
46	71.2	5.7	1662	5	PCT-US91-06399-5	Sequence 5, Appl
47	66.4	5.3	1641	1	US-08-705-771-9	Sequence 1, Appl
48	65.4	5.2	2233	1	US-08-496-641-1	Sequence 1, Appl
49	65.2	5.2	1952	1	US-08-333-358-1	Sequence 1, Appl
50	65.2	5.2	1952	1	US-08-463-694-1	Sequence 1, Appl
51	65.2	5.2	1952	1	US-08-463-694-1	Sequence 1, Appl
52	64.8	5.2	2095	1	US-08-694-501-1	Sequence 1, Appl
53	64.8	5.2	2095	1	US-08-694-501-1	Sequence 1, Appl
54	64.8	5.2	2095	1	US-08-694-501-1	Sequence 1, Appl
55	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
56	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
57	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
58	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
59	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
60	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
61	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
62	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
63	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
64	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
65	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
66	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
67	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
68	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
69	64	5.1	1245	2	US-08-891-298-2	Sequence 2, Appl
70	63.6	5.1	2094	4	US-09-106-194-1	Sequence 1, Appl
71	62	4.9	2130	1	US-07-952-800-1	Sequence 1, Appl
72	61.2	4.9	2468	1	US-08-333-358-11	Sequence 1, Appl
73	61.2	4.9	2468	1	US-08-333-358-11	Sequence 1, Appl
74	61.2	4.9	2468	1	US-08-333-358-11	Sequence 1, Appl
75	60.8	4.8	1746	4	US-08-694-501-11	Sequence 1, Appl
76	60.8	4.8	2005	2	US-08-484-200-1	Sequence 1, Appl
77	60.8	4.8	2005	2	US-08-484-200-1	Sequence 1, Appl
78	60.8	4.8	2005	2	US-08-484-200-1	Sequence 1, Appl
79	60.8	4.8	2005	2	US-08-484-200-1	Sequence 1, Appl
80	60.8	4.8	2005	2	US-08-484-200-1	Sequence 1, Appl
81	60.4	4.8	1866	4	US-08-336-408B-1	Sequence 1, Appl
82	59.2	4.7	2234	1	US-07-952-800-1	Sequence 1, Appl
83	59.2	4.7	2234	1	US-07-952-800-1	Sequence 1, Appl
84	59.2	4.7	2234	1	US-07-952-800-1	Sequence 1, Appl
85	59.2	4.7	2234	1	US-07-952-800-1	Sequence 1, Appl
86	59	4.7	1662	5	PCT-US91-00399-7	Sequence 7, Appl
87	59	4.7	1662	5	PCT-US91-00399-7	Sequence 7, Appl
88	59	4.7	1662	5	PCT-US91-00399-7	Sequence 7, Appl
89	58.8	4.7	1787	2	US-08-372-652-14	Sequence 14, Appl
90	58.8	4.7	1787	2	US-08-372-652-14	Sequence 14, Appl
91	58.8	4.7	1787	2	US-08-372-652-14	Sequence 14, Appl
92	56.8	4.5	1344	4	US-08-553-648A-9	Sequence 9, Appl
93	56.8	4.5	1344	4	US-08-553-648A-9	Sequence 9, Appl
94	56.8	4.5	1344	4	US-08-553-648A-9	Sequence 9, Appl
95	56.8	4.5	1344	4	US-08-553-648A-9	Sequence 9, Appl
96	56.8	4.5	1344	4	US-08-553-648A-9	Sequence 9, Appl
97	54.4	4.3	1514	4	US-08-653-648A-4	Sequence 4, Appl
98	54.4	4.3	1514	4	US-08-653-648A-4	Sequence 4, Appl
99	54.4	4.3	1514	4	US-08-653-648A-4	Sequence 4, Appl
100	54.4	4.3	1514	4	US-08-653-648A-4	Sequence 4, Appl

[illegible]

```

01 008 GTCACAGCCACCAACCAAGCTGGACAGACGACGGGCGAAGATGCGAGTTCGCTTCAGCGACCGG 114
01 1294 GTCAACGACCTCAAGATCTGCACAGACGCGCGACAGCTGATGATTTGGTGAACGATGGG 1458
01 1141 ATCTAAGCTTTGGTTGGATGATATGCTCAACAGACGATATCTCTCCACCAACCAATGAC 1200
01 1454 AGGATGATCTTGATTTGGTGATATGCTCAACAGACGATATCTCTCCACCAACCAATGAC 1418
01
01 1201 GAGTGGCTTAACCTCTGATGATCTCTGCTGCTCAATGATAGCATGCGAG 1247
01 1414 GAGCTGACCAACCTGCTGATGATCTCTGCTGCTCAATGATAGCATGAG 1460
01
01 RESULT 4
01 05-08-846-620A-1
01 Sequence 1, Application US/08836620A
01 Patent No. 5958710
01 GENERAL INFORMATION:
01 APPLICANT:
01 TITLE OF INVENTION: orphan receptor
01 NUMBER OF SEQUENCES: 19
01 COMPUTER READABLE FORM:
01 MEDIUM TYPE: floppy disk
01 COMPUTER: IBM PC compatible
01 OPERATING SYSTEM: PC-DOS/MS-DOS
01 SOFTWARE: Patent to Release #1.0, Version #1.25 (BPO)
01 CURRENT APPLICATION DATA:
01 APPLICATION NUMBER: US/08/846-620A
01 FILING DATE:
01 PRIOR APPLICATION DATA:
01 APPLICATION NUMBER: PCT/EP96/03933
01 FILING DATE:
01 APPLICATION NUMBER: GB 9518272.1
01 FILING DATE: 08-SEP-1995
01 PRIOR APPLICATION DATA:
01 APPLICATION NUMBER: GB 9605550.4
01 FILING DATE: 15-MAR-1996
01 PRIOR APPLICATION DATA:
01 APPLICATION NUMBER: GB 9607532.0
01 FILING DATE: 11-APR-1996
01 PRIOR APPLICATION DATA:
01 APPLICATION NUMBER: GB 9609570.5
01 FILING DATE: 08-MAY-1996
01 INFORMATION FOR SEQ ID NO: 1:
01 SEQUENCE CHARACTERISTICS:
01 LENGTH: 2568 base pairs
01 TYPE: nucleic acid
01 STRANDEDNESS: double
01 TOPOLOGY: linear
01 ORIGINAL SOURCE:
01 ORGANISM: Rattus rattus
01 US-08-846-620A-1
01
01 Query Match 74.5%; Score 937; OR 2; Length 2568;
01 Best Local Similarity 84.7%; Prod. No. 76.249;
01 Matches 1051; Conservative 0; Mismatches 190; Indels 0; Gaps 0;
01
01 1 ATGATATTACGATTGCTGACGATGCTGATCTTCGACGACGCGGCTGAGTACAC 60
01 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 448 ATTAAATTACAGCTTCTGCTGACGACGATGACGATGACGATGACGATGACG 507
01 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 61 ACAAACCCAAATGCTCTGCTGACGACGATGACGATGACGATGACGATGACG 120
01 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 508 CAAAGCTCAAGATGCTGCTGACGACGATGACGATGACGATGACGATGACG 567
01 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 121 GATTATATGATCTGATGCTGACGATGACGATGACGATGACGATGACGATG 180
01 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 568 CAATGATGCTCTGATGACGACGATGACGACGATGACGACGATGACGATG 627
01 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 181 GAAATACCTTATCTGTAAGATGACGACGATGACGACGATGACGACGATGACG 240
01 1111 1111 1111 1111 1111 1111 1111 1111 1111
01 628 GATCAATCTTACTGCTGAAGATGACGACGATGACGACGATGACGACGATG 687

```



```

1  SWAPPING SYSTEM:  PG-PPS/XS-PPS
2  SOFTWARE:  ASCII text
3  CURRENT APPLICATION DATA:
4  APPLICATION NUMBER:  05/08/897,719
5  FILING DATE:
6  CLASSIFICATION:
7  PRIOR APPLICATION DATA:
8  APPLICATION NUMBER:  08/260,452
9  FILING DATE:
10 APPLICATION NUMBER:  08/076,327
11 FILING DATE:  14-JUN-1993
12 ATTORNEY/AGENT INFORMATION:
13 NAME:  Giulio A. Decanti, Jr.
14 REGISTRATION NUMBER:  31,503
15 REFERENCE/DOCKET NUMBER:  RAI 018CP
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE:  (617) 227-7400
18 TELEFAX:  (617) 227-5941
19 INFORMATION FOR SEQ ID NO:  9:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH:  4963 base pairs
22 TYPE:  nucleic acid
23 STRANDEDNESS:  double
24 TOPOLOGY:  circular
25 MOLECULE TYPE:  DNA (genomic)
26 ORIGINAL SOURCE:
27 ORGANISM:  Human cytomegalovirus
28 IMMEDIATE SOURCE:
29 CLONE:  pUHD 1634
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
98
```



```

1  APPLICANT: Gossen, Manfred
2  APPLICANT: Bujard, Hermann
3  APPLICANT: Salfeld, Jochem
4  APPLICANT: Voss, Jeffrey
5  TITLE OF INVENTION: Animal Transgenic for a Tetracycline-Controlled
6  NUMBER OF SEQUENCES: 10
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: Lathive & Cockfield
9  STREET: 60 State Street, Suite 510
10 CITY: Boston
11 STATE: Massachusetts
12 COUNTRY: USA
13 ZIP: 02109-1875
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: ASCII text
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/481,970
21 FILING DATE:
22 CLASSIFICATION: 800
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 08/260,452
25 FILING DATE: 14-JUN-1994
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: 08/076,427
28 FILING DATE: 14-JUN-1993
29 ATTORNEY/AGENT INFORMATION:
30 NAME: Giulio A. DeConti, Jr.
31 REGISTRATION NUMBER: 31,503
32 REFERENCE/PACKET NUMBER: RRI-013CP2
33 TELECOMMUNICATION INFORMATION:
34 TELEPHONE: (617) 227-7400
35 TELEFAX: (617) 227-5941
36 INFORMATION FOR SEQ ID NO: 8:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 6244 base pairs
39 TYPE: nucleic acid
40 STRANDEDNESS: double
41 TOPOLOGY: circular
42 MOLECULE TYPE: DNA (genomic)
43 ORIGINAL SOURCE:
44 ORGANISM: Human cytomegalovirus
45 STRAIN: Towne (hcmv)
46 IMMEDIATE SOURCE:
47 CLONE: PUHD BGR3
48 US-08-481-970-8

```

	Query Match	7.1%	Score 89	DB 2:	Length 6244	
	Post Local Similarity	66.38	Pred No. 5.5e-14			
	Matches 129	Conservative	0	Mismatches 45	Indels	0
	Caps	0				
OY	293 TCTGCAATGATTAAATACCATATATATAGATTAATTTTATATATGAAAGCATAAAT	45.2				
D6						
OY	2365 TCTGTGGATGAAGAATATGATATCTGTGATATATATGATATGATATGATATGATAT	24.2				
D6						
OY	453 CTTTTTAAAGCAATATTAAGGACATATATGATATATTTTCTATATTAATATAT	41.2				
D6						
OY	2425 TTCTCTTTAAAAGGCATATTAAGGATATTAATATTTATATGTCGAAATAAATGAT	24.8				
D6						
OY	413 GTCAAATCGATAAAAAAGGCTTAAGATATATGATATGATATGATATGATATGATAT	47.2				
D6						
OY	2485 GCATTTTGATATAAAATCCGATATAAAAATTTGGCGGCGGTTCGGCTTAAATATGCTCT	25.4				
D6						
OY	473 AACTGGAAATGGT	48.5				
D6						
OY	2545 AACGTGCATGGT	25.7				
D6						

RESULT 22
US-08-897-719-8
Sequence B, Application US/080847719

```

1 Patent No.: 922927
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Gossen, Manfred
6
7 APPLICANT: Bugard, Hermann
8
9 APPLICANT: Salfeld, Jochem
10
11 APPLICANT: Voss, Jeffrey
12
13 TITLE OF INVENTION: Tight Control of Gene Expression in Eucaryotic
14
15 TITLE OF INVENTION: Cells by Tetracycline-responsive Promoters
16
17 NUMBER OF SEQUENCES: 10
18
19 CORRESPONDENCE ADDRESS:
20
21 ADDRESSEE: Lahive & Cockfield
22
23 STREET: 60 State Street
24
25 CITY: Boston
26
27 STATE: Massachusetts
28
29 COUNTRY: USA
30
31 ZIP: 02109-1875
32
33 COMPUTER READABLE FORM:
34
35 MEDIUM TYPE: Floppy disk
36
37 COMPUTER: IBM PC compatible
38
39 OPERATING SYSTEM: PC-DOS/MS-DOS
40
41 SOFTWARE: ASCII text
42
43 CURRENT APPLICATION DATA:
44
45 APPLICATION NUMBER: US/08/897,719
46
47 FILING DATE:
48
49 CLASSIFICATION:
50
51 PRIOR APPLICATION DATA:
52
53 APPLICATION NUMBER: 08/260,452
54
55 FILING DATE:
56
57 APPLICATION NUMBER: 08/076,327
58
59 FILING DATE: 14-JUN-1994
60
61 ATTORNEY/AGENT INFORMATION:
62
63 NAME: Giulio A. Decortis, Jr.
64
65 REGISTRATION NUMBER: 31,503
66
67 REFERENCE/DOCKET NUMBER: BRL-013CP
68
69 TELECOMMUNICATION INFORMATION:
70
71 TELEPHONE: (617) 227-7400
72
73 TELEFAX: (617) 227-5941
74
75 INFORMATION FOR SEQ ID NO: 8:
76
77 SEQUENCE CHARACTERISTICS:
78
79 LENGTH: 6244 base pairs
80
81 TYPE: nucleic acid
82
83 STRANDEDNESS: double
84
85 TOPOLOGY: circular
86
87 MOLECULE TYPE: DNA (genomic)
88
89 ORIGINAL SOURCE:
90
91 ORGANISM: Human; cytomegalovirus
92
93 STRAIN: Towne (HCMV)
94
95 IMMEDIATE SOURCE:
96
97 CLONE: pUHD BGR3
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
94
```

Query Matched	7.1%	Score 89	PI 2	Length 6244
Best Local Similarity	66.3%	Pred. No. 5	Se-14	
Matches 128	Conservative	0	Mismatches 65	Indels 0
			Gaps 0	
QY	294 TCTGGGACATTAAGGATATGGAGATACAGATAGAGATGAGAGATATGAAG	352		
DB	111 111 1111 111 1111 111 1111 1111 1111 111111			
	2465 TCTGGGACATTAAGGATATGGAGATACAGATAGAGATGAGAGATATGAAG	2424		
QY	353 GCTTTTAAAGAAATATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	412		
DB	111 111111111 111111 111 1111 1111 111 1111 111			
	2425 TCTTTTAAAGAAATATCAATCAATCAATCAATCAATCAATCAATCAAT	2404		
QY	413 GTACAAATCAATAAAGAAATGAGAGATGAGAGATGAGAGATGAGAGAT	472		
DB	111 111111 111 111 1111 111 111 1111 111 1111 111			
	2485 GTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2544		
QY	473 AAGTGGGAATGGT	485		
DB	111 111111			
	2545 AAGCTGGGAATGGT	2557		

RESULTAT 23


```

1 CLASSIFICATION: 415
2
3 ATTORNEY/AGENT INFORMATION:
4 NAME: KIT, Gordon
5 REGISTRATION NUMBER: 40,764
6 REFERENCE/DOCKET NUMBER: A-6612
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (202) 293-7060
9 TELEFAX: (202) 293-7060
10 INFORMATION FOR SEQ ID NO: 1:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 4014 base pairs
13 TYPE: nucleic acid
14 STRANDEDNESS: single
15 TOPOLOGY: linear
16 MOLECULE TYPE: cDNA
17 HYPOTHEICAL: NO
18
19 US-08-629-939-1
20
21 Query Match 6-4% Score 81; DB 1; Length 4014;
22 Best Local Similarity 63.7%; Pred. No. 5,4e-12;
23 Matches 12; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
```

```

1 LENGTH: 3014 base pairs
2 TYPE: nucleic acid
3 STRANDEDNESS: single
4 TOPOLOGY: linear
5 MOLECULE TYPE: cDNA
6 HYPOTHEetical: NO
7
8 MS: OR 759-874 1

```

Query Match: 6.4%; Score 81; DB 1; Length 1014;
 Best Local Similarity: 63.7%; Prod. No. 5,40512;
 Matches: 124; Conservative: 0; Mismatches: 70; Indels

```

APPLICATION NUMBER: 1HS/07/134, 140
FILING DATE: 17 DEC-1987
PRIORITY DATA:
P NO: 1
LENGTH: 2989
6-1
Y MATCH: 6.2% SCORE 7
LOCAL SIMILARITY: 6.19% PROD. N
THS 140: CONSERVATIVE 0; MISM

```

Query Match: 6.28; Score 78; DB 6; Length 2089;
 Host Local Similarity: 61.98;
 Prod. No. 3.4e+11;
 Matches: 140; Conservative: 0; Mismatches: 85; Indels: 1; Caps: 1

[illegible][illegible]

27	474	AACTGGGAAATGGT	485
106	2061	AGGCTGGGATGGT	2079

Db 740 AGTGTTCGAGTGGGAATGTTC AAAGAAATCTCTCAGGAAATGAAAG 789

REFSHEET 28

1. Entry No. 5260432
 2. APPLICANT: TAKAKU, FUMIMARO; ISHIKAWA, TARASHI; IMAMOTO, MICHIO;
 3. EVANS, RONALD M.; JEMISON, ROZILLO
 4. TITLE OF INVENTION: HUMAN GAMMA GELINOLIC ACID RECEPTOR DNA
 5. NUMBER OF SPECIMENS: 4
 6. CURRENT APPLICATION NUMBER: 08/073,407
 7. APPLICATION NUMBER: 08/073,407
 8. FILING DATE: 22 JUN 1989
 9. SEQ. ID. NO.: 1
 10. LENGTH: 1576
 11. 5260432.1

1 RESULT 30
 2 US-09-277-078-1
 3 Sequence 1, Application US/09277078
 4 Patent No. 6,612,549
 5 GENERAL INFORMATION:
 6 APPLICANT: Sakurada, Kazuhiro
 7 APPLICANT: Palmer, Theo
 8 APPLICANT: Gage, Fred H.
 9 TITLE OF INVENTION: REGULATION OF TYROSINE HYDROXYLASE
 10 FILE REFERENCE: 07257/031001
 11 CURRENT APPLICATION NUMBER: 05/09,277, 078
 12 CURRENT FILING DATE: 1999-04-26
 13 NUMBER OF SEQ ID NOS: 60

Quercy March 6.4M; Score 80.8; DH 6; Length 1576;

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 4.0
; SEQ ID NO 1

```

Matches 127; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

TYPE: DNA

286 TCGTCTTCAATACCAATCCGATATCACTATACGAGCTCTGATACGGA 345

NAME/KEY: _____

497 195.110.161AA16AA6101010601A0CA1A1660678A6310101616AA635 526

```

; LAXATION: (1) .. (900)
HC=10+077-078+1

```

396 101AAAGGAGAGGATCTCAAGGAGAAATAAATCAATATCTGCTTCAAGGATCA 403

(More Match) 5 08: Score 74 4: 100th 1797:

[illegible]

Best local similarity	37.98;	Prod. NO. 2,000-107
Matches	137;	Conservative
	0;	Mismatches
	96;	Indels
	0;	idps

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

284 TTTGGGCTCTGACCGATTACCGATCGGGATATCACTATGGAGTCTGGTCTTTCAAC 344

10b 785 TGTGGCTTTTGGGTGACAAAGGGCTGTACAGCAATACGGTGTTCGCACTTGTGAGG 844

344 GATGTTAAGGAGCTTTTAAAAAGAAAGATTTTAAAGTACATTAATGATTTTATTTTGTCTCAAGTAA 405

845 GTGCAAGCTTCTTTAAGCGACGGTGAATAAAAAAATCTTAATATGTGTCTTACCA 904

REVISION 29

404 CAAATCAGTGTAAATCGATAAAAAACCGCGGGAAGAGCTGCGAGGCGCTTCCGGA 463

Patent No. 5,223,606

905 ATAAAAATTCGGCACTGATACGCGCGAAATCTTCTGAGTA¹CTCTGGT¹¹TA²AA³A 964

PERKINS, DEJHAN, ANNE

464 AGTGTACCGAAGTCGTTAACTGTGGCTCCCGAGAGAGAGAT 511

PROTEIN INAPPROPRIATELY EXPRESSED IN HUMAN HEPATOCELLULAR CARCINOMA

100 965 AGTGGCTGGCTGTTGGGATGGTTAAAGCAAGTGGTTCCGCAAGCAAT 1012

```
CHICKENNT ALPLIAT' ATION DATA;
```

RESULT 31

US-08-592-483-3
Sequence 3, Application US/08592383
Patent No. 5830760
GENERAL INFORMATION:
APPLICANT: Tsai, S. and S.J. Collins
TITLE OF INVENTION: "Hematopoietic Cell Lines Bearing Altered Polynucleotide Receptor"
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
STEPPED: 9800 Pacific First Centre, 1420 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-5.25 inch, 1.2MB storage
COMPUTER: IBM PC/486 Compatible
OPERATING SYSTEM: MS-DOS 4.01
SOFTWARE: Word for Windows 5.01-1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,383
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/099,242
FILING DATE: July 28, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Broderick, Thomas F.
REGISTRATION NUMBER: 31,332
REFERENCE/BOOK NUMBER: FHP-1-7190
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-206-692-9100; 1-206-224-0709(direct)
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2658 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: page 4, RAR-alpha403 dominant negative; deleted of 1311-1546 of SEQ.
IMMEDIATE SOURCE:
LIBRARY: cDNA
US-08-592-483-3

Query Match

5.98; Score 74.4; DB 2; Length 2658;
Best Local Similarity 60.38; Pred. No. 2 9e-19;
Matches 123; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 286 TGGCGCTTCTAGCGATTACGATGAGATATCATAGAGCTGGTCGTGAGAGCA 345
DB 364 TCGTTGTCTGTGAGGACAGAGTGTAGGCTACGCTAGGCTGAGAGGTC 423
QY 446 TGTAAAGCGCTTTTAAAGAACATTCAGAGCATATGATATATTGTCACCTACA 405
DB 424 TCCAGAGCGCTTCTCGCCGACAGCAAGAACATGCTGTACACGTCACACGAGC 483
QY 406 AATCAGCTGATCATATCAATAAAGACGAGCAGAGATGTCAGGCTGCGGACATGGAGAG 465
DB 484 AAGAACTGATCATATCAACACAAAGTGAACCGAATGCTGCACTATGTCAGAGAG 543
QY 466 TGTACGAAGTGGGAATGCTGAAG 489
DB 544 TCGTTGAAGTGGGCATGTCAGAG 567

RESULT 32

US-08-095-728B-3
Sequence 3, Application US/08095728B
Patent No. 5843642
GENERAL INFORMATION:
APPLICANT: DMITROVSKY, ETHAN
APPLICANT: WARRELL JR, RAYMOND P

APPLICANT: MILLER JR, WILSON H
APPLICANT: FRANKEL, STANLEY
TITLE OF INVENTION: METHODS FOR THE DETECTION AND
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 AVENUE OF THE AMERICANS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patcutild Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,728B
FILING DATE: 21-JUL-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/673,838
FILING DATE: 22-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN P.
REGISTRATION NUMBER: 28,678
REFERENCE/BOOK NUMBER: 38694-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525

INFORMATION FOR SEQ. ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 2928 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: hRAR ALPHA
FEATURES:

NAME/KEY: CDS
LOCATION: 103..1488
OTHER INFORMATION:
US-08-095-728B-3

Query Match

5.98; Score 74.4; DB 2; Length 2928;
Best Local Similarity 60.38; Pred. No. 3e-10;
Matches 123; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 286 TGGCGCTTCTAGCGATTACGATGAGATATCATAGAGCTGGTCGTGAGAGCA 345
DB 364 TCGTTGTCTGTGAGGACAGAGTGTAGGCTACGCTAGGCTGAGAGGTC 423
QY 346 TGTAAAGCGCTTTTAAAGAACATTCAGAGCATATGATATATTGTCACCTACA 405
DB 424 TCCAGAGCGCTTCTCGCCGACAGCAAGAACATGCTGTACACGTCACACGAGC 483
QY 406 AATCAGCTGATCATATCAATAAAGACGAGCAGAGATGTCAGGCTGCGGACATGGAGAG 465
DB 484 AAGAACTGATCATATCAACACAAAGTGAACCGAATGCTGCACTATGTCAGAGAG 543
QY 466 TGTACGAAGTGGGAATGCTGAAG 489
DB 544 TCGTTGAAGTGGGCATGTCAGAG 567

RESULT 33

US-08-095-02320A-3
Sequence 3, Application US/0809502320A
Patent No. 5843642
GENERAL INFORMATION:
APPLICANT: Sloan-Kettering Institute, For Cancer Research
TITLE OF INVENTION: METHODS FOR DETECTION AND TREATMENT OF CANCER

OTHER INFO
US-08-095-728B-1

[illegible]

```

1 RESULT 49
2 US-008 412 412A-1/0
3 Sequence 1, Application 05/08 412412A
4 Patent No. 5550024
5 GENERAL INFORMATION:
6 APPLICANT: Rothschild, Max F.
7 APPLICANT: Tugde, Christopher K.
8 APPLICANT: Jacobson, Carol D.
9 APPLICANT: Miloham, Alan J.
10 APPLICANT: Plashow, Graham S.
11 APPLICANT: Vaske, David A.
12 TITLE OF INVENTION: Genetic Markers for Pig Litter Size
13 NUMBER OF SEQUENCES: 9
14 CORRESPONDENCE ADDRESS:
15 ADDRESSEE: Zaffey, McKee, Thonley, Voorhees & Seaso
16 STREET: 801 Grand Suite 3200
17 CITY: Des Moines
18 STATE: Iowa
19 COUNTRY: United States
20 ZIP: 50309
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: Floppy disk
23 COMPUTER: IBM PC compatible
24 OPERATING SYSTEM: PC-DOS/MS-DOS
25 SOFTWARE: Patent Release #1.0, Version #1.00
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: 05/08/412, 412A
28 FILING DATE: 26-SEP-1994
29 CLASSIFICATION: 445
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: US 07/961,819
32 FILING DATE: 19-007-1992
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: US 07/687,708
35 FILING DATE: 19-APR-1991
36 ATTORNEY/AGENT INFORMATION:
37 NAME: Nobel, Heidi S.
38 REGISTRATION NUMBER: 47,719
39 REFERENCE/DOCKET NUMBER: BRUC-4
40 TELECOMMUNICATION INFORMATION:
41 TELEPHONE: 515-288-4667
42 TELEFAX: 515-288-1438
43 INFORMATION FOR SEQ ID NO: 1:
44 SEQUENCE CHARACTERISTICS:
45 LENGTH: 4460 base pairs
46 TYPE: nucleic acid
47 STRANDEDNESS: double
48 TOPOLOGY: linear
49 MOLECULE TYPE: DNA (genomic)
50 HYPOTHEetical: NO
51 ANTI-SENSE: NO

```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 303..304
OTHER INFORMATION: /note: "Gap in DNA sequence of
OTHER INFORMATION: undetermined length"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1197..1198
OTHER INFORMATION: /note: "Gap in DNA sequence of
OTHER INFORMATION: undetermined length"
FEATURE:
NAME/KEY: exon
LOCATION: 534..600
US-08-312-412A-1

Query Match      5 94; Score 74.4; DB 1; Length 3466;
Best Local Similarity 81.7%; Pred. No. 3,2e-10;
Matches 98; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 475 AGGACATATGATTAATTTGTCAGCTACCAATCAGTACATGCAATAAAAAGCCGCG 434
DB 602 AATGATATATGATTAATTTGTCAGCTACCAATCAGTACATGCAATAAAAAGCCGCG 543
QY 435 CAAGAGCTGCGAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCT 434
DB 542 GAAAGCTTTTGA-TGTTGCTGCTTAAATGATGATGATGATGATGATGATGATGATGAT 444

RESULT 40
US-08-892-747-13
Sequence 14, Application US/08892747
Patent No. 6057153
GENERAL INFORMATION:
APPLICANT: Shaji T. George, Michael M. Martina Werner,
APPLICANT: Umberto Pace and Allan R. Goldberg
TITLE OF INVENTION: Stabilized External Guide Sequences
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patricia L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/892,747
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/472,556
FILING DATE: January 13, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/00514
FILING DATE: January 19, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patricia L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: 111109C102
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 3511 base pairs
TYPE: nucleic acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..3511
OTHER INFORMATION: /function: "PMU-RAR' DNA Sequence."
US-08-892-747-13

Query Match      5 94; Score 74.4; DB 3; Length 3511;
Best Local Similarity 60.3%; Pred. No. 3,2e-10;
Matches 123; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 386 TCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345
DB 1898 TCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1867
QY 346 TGTAAAGCCCTTTTAAAGAGCATTCAGGACATATGATTAATTTGTCAGCTACA 405
DB 1888 TGTAAAGCCCTTTTAAAGAGCATTCAGGACATATGATTAATTTGTCAGCTACA 1927
QY 406 AATGATATATGATTAATTTGTCAGCTACCAATCAGTACATGCAATAAAAAGCCGCG 405
DB 1928 AATGATATATGATTAATTTGTCAGCTACCAATCAGTACATGCAATAAAAAGCCGCG 1987
QY 466 TGTAAAGCCCTTTTAAAGAGCATTCAGGACATATGATTAATTTGTCAGCTACA 489
DB 1988 TGTAAAGCCCTTTTAAAGAGCATTCAGGACATATGATTAATTTGTCAGCTACA 2011

```

Search completed: June 16, 2003, 04:20:50
 Job time : 62.1909 secs


```

UY 608      GGAAGCTGGTCTGCTGAGAGCTGAGAGAGAGTAACTGATCTGCTG 657
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2782  GCTCTTAAGAAAGAAAGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2841
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 658  CTTGAGAGCTTGAAGCTGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 714
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2842  TTTGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2901
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 715  GCTCTTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2902  GCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2961
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 775  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2962  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3021
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 835  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3022  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3081
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 895  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3082  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3141
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 955  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3142  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3201
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1015  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3202  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3261
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1075  GCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1134
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3262  GCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3321
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1132  AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3322  AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3381
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1192  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1247
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3382  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3447
              ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

UY 271  AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538  AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 431  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 490
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 598  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 491  ATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 658  ATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 451  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 502
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 718  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 503  ---GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 778  AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 559  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 599
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 838  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 600  ---GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 898  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 658  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 958  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 715  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1018  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 775  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1078  TGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 835  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1138  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 895  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1198  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 955  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1258  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1015  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1318  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1075  GCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1134
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1378  GCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1132  AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1438  AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
              ||| ||| ||| ||| ||| ||| ||| ||| |||
UY 1192  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1247
              ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1498  GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1554
              ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 16
 US-10 198-846-11405
 Sequence 11405, Application US/10198846


```

APPLICANT: Penn, Sharon G.
APPLICANT: Kark, David R.
APPLICANT: Haezel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DEPLETED SINCE EXON NUCLEIC ACID PROBES USED FOR
FILE REFERENCE: Aeonica-X-1
CURRENT APPLICATION NUMBER: US-09-864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/642,366
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 24/263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/000666
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-10
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Mosaic Sequence Listing Engine vers. 1.1
SEQ ID NO 32872
LENGTH: 300
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL161756.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.46
OTHER INFORMATION: SWISSPROT HIT: g92731, FVALUE 6.0e-02
OTHER INFORMATION: EST_HUMAN HIT: BE74783.1, FVALUE 6.90e-02
OTHER INFORMATION: NI HIT: Ab006500.1, FVALUE 0.00e+00
US-09-864-761-32872

Query Match: 23.9%; Score 300; DB 10; Length 300;
Best Local Similarity 100.0%; Pred. No. 40-83;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 121 TGGTGGTGGAGCGGCTGAGAGCGGAGAGTATGTCATGCTGCTGGAGGTTGGAGCGG 180
GY 674 CCGATGCTGATGAGCGGCGGAGTGGGCTTATACGAGGCGCTGATGATGATGCG 733
DB 181 GCGATGCTGATGAGCGGCGGAGTGGGCTTATACGAGGCGCTGATGATGATGCG 240
GY 734 TGGTGGTGGAGCGGCTGAGAGCGGAGTATGTCATGCTGCTGGAGGTTGGAGCGG 793
DB 241 TGGTGGTGGAGCGGCTGAGAGCGGAGTATGTCATGCTGCTGGAGGTTGGAGCGG 300

RESULT 20
US-09-918-995-15134
Sequence 15134, Application us/09918995
Publication No. US2004007362A1
GENERAL INFORMATION:
APPLICANT: Hysq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 15134
LENGTH: 451
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc:feature
LOCATION: (1)-(451)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-15134

Query Match: 22.1%; Score 278; DB 9; Length 451;
Best Local Similarity 90.2%; Pred. No. 4-60-76;
Matches 296; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```



```

1  PEPID APPLICATION NUMBER: PEPID05017090601
2  PRIOR FILING DATE: 2001-01-30
3  PEPID APPLICATION NUMBER: PEPID05017090670
4  PRIOR FILING DATE: 2001-01-30
5  PRIOR APPLICATION NUMBER: US 60/244,687
6  PRIOR FILING DATE: 2000-09-21
7  PEPID APPLICATION NUMBER: US 09/608,408
8  PRIOR FILING DATE: 2000-06-30
9  PRIOR APPLICATION NUMBER: US 09/774,203
10 PRIOR FILING DATE: 2001-01-29
11 NUMBER OF SEQ ID NOS: 49117
12 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
13 SEQ ID NO 24607
14 LENGTH: 306
15 TYPE: DNA
16 ORGANISM: Homo sapiens
17 FEATURE:
18 OTHER INFORMATION: MAP TO AL161756.1
19 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6
20 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.8
21 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
22 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.2
23 OTHER INFORMATION: EXPRESSED IN KIDNEY, SIGNAL = 4.1
24 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.7
25 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6
26 OTHER INFORMATION: SWISSPROT HIT: Q92731, EVALU: 9.00e-58
27 OTHER INFORMATION: EST HUMAN HIT: BF10841.1, EVALU: 1.00e-46
28 OTHER INFORMATION: NT HIT: AB006590.1, EVALU: 0.00e+00
29 US-09-864-761-24607

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```
2 PRIOR APPLICATION NUMBER: US 60/2246,459
3 PRIOR FILING DATE: 2000 09 27
4 PRIOR APPLICATION NUMBER: PCT/US01/00666
5 PRIOR FILING DATE: 2001 01 40
6 PRIOR APPLICATION NUMBER: PCT/US01/00667
7 PRIOR FILING DATE: 2001 01 40
8 PRIOR APPLICATION NUMBER: PCT/US01/00664
9 PRIOR FILING DATE: 2001 01 40
10 PRIOR APPLICATION NUMBER: PCT/US01/00669
11 PRIOR FILING DATE: 2001 01 40
12 PRIOR APPLICATION NUMBER: PCT/US01/00665
13 PRIOR FILING DATE: 2001 01 40
14 PRIOR APPLICATION NUMBER: PCT/US01/00668
15 PRIOR FILING DATE: 2001 01 40
16 PRIOR APPLICATION NUMBER: PCT/US01/00664
17 PRIOR FILING DATE: 2001 01 40
18 PRIOR APPLICATION NUMBER: PCT/US01/00662
19 PRIOR FILING DATE: 2001 01 40
20 PRIOR APPLICATION NUMBER: PCT/US01/00661
21 PRIOR FILING DATE: 2001 01 40
22 PRIOR APPLICATION NUMBER: PCT/US01/00670
23 PRIOR FILING DATE: 2001 01 40
24 PRIOR APPLICATION NUMBER: US 60/234,687
25 PRIOR FILING DATE: 2000 09 21
26 PRIOR APPLICATION NUMBER: US 09/608,408
27 PRIOR FILING DATE: 2000 06 30
28 PRIOR APPLICATION NUMBER: US 09/774,203
29 PRIOR FILING DATE: 2001 01 29
30 NUMBER OF SEQ ID NOS: 49117
31 SOFTWARE: Ambis Sequence List and Edit v6.1.1
32 SEQ ID NO 7903
33 LENGTH: 597
34 TYPE: DNA
35 ORGANISM: Homo sapiens
36 FEATURE:
37 OTHER INFORMATION: MAP TO AL161756.1
38 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 6
39 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 5.8
40 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 4.3
41 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 5.2
42 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 4.1
43 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 4.7
44 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 6
45 US 09 864 761 7903
46
47 Query Match 16.08: Score 204, Hs 109, Length 597:
48 Best Local Similarity 100.08: Prod. No. 9,6e 53:
49 Matches 204: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
50
51 1 ATGAAATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGCGGCAATG 60
52 |||||||
53 252 ATGAAATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGGCAATG 120
54 |||||||
55 412 ATGAAATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGGCAATG 180
56 |||||||
57 121 GATTTATGCAATGCTGATGCGGAAGCTTCAAAAGAGCTGCGGCTCAAGATGCGTA 180
58 |||||||
59 472 GATTTATGCAATGCTGATGCGGAAGCTTCAAAAGAGCTGCGGCTCAAGATGCGTA 431
60 |||||||
61 181 GAAATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGGCAATG 204
62 |||||||
63 432 GAAATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGGCAATG 454
64 |||||||
65
66 RESULT 29
67 US 09 747 255 4
68 Sequence 29: Application US/09747255
69 Patent No. US20020073441A1
70 GENERAL INFORMATION:
71 APPLICANT: Biomedical, Almazaz
72 APPLICANT: Rossi, Brian D.
73 FILE REFERENCE: TSKI 725.1
```

```
1 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
2 TITLE OF INVENTION: PROTEOLYTIC ACTIVITY
3 FILE REFERENCE: 11203-003001
4 CURRENT APPLICATION NUMBER: US/09/737,255
5 CURRENT FILING DATE: 2000-12-13
6 NUMBER OF SEQ ID NOS: 6
7 SOFTWARE: FASTSP for Windows Version 4.0
8 SEQ ID NO 4
9 LENGTH: 973
10 TYPE: DNA
11 ORGANISM: Mus musculus
12 FEATURE:
13 NAME: 725
14 LOCATION: (1)...(973)
15 US 09 747 255-4
16
17 Query Match 16.08: Score 200.8: Hs 109, Length 973:
18 Best Local Similarity 61.78: Prod. No. 6,1e 52:
19 Matches 438: Conservative 0: Mismatches 207: Indels 4: Gaps 1:
20
21 703 CCTTTCAGCAAGCTTCATGAGGATGCTGCAATGCAATGCAAGGCGGCGGCA 762
22 |||||||
23 184 CCTTTCAGCAAGCTTCATGAGGATGCTGCAATGCAATGCAAGGCGGCGGCA 243
24 |||||||
25 763 CATATGATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGGCA 822
26 |||||||
27 244 CATATGATTAAGCTTCGACAAATGCAATGCAATGCAAGGCGGCGGCA 903
28 |||||||
29 823 GTCGAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGACAA 882
30 |||||||
31 304 GTCGAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGACAA 363
32 |||||||
33 883 ATTGATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 942
34 |||||||
35 364 ATTGATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 423
36 |||||||
37 943 AATGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGACAAAT 1002
38 |||||||
39 424 AATGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGACAAAT 483
40 |||||||
41 1003 GATATGATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTC 1062
42 |||||||
43 484 GATATGATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTC 543
44 |||||||
45 1063 TCTGATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTC 1122
46 |||||||
47 544 TCTGATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 603
48 |||||||
49 1123 GATATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 1179
50 |||||||
51 604 GATATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 664
52 |||||||
53 1240 GATATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 1299
54 |||||||
55 724 GATATTAAGCTTCGACAAATGCTTCGACAAATGCTTCGACAAATGCTTCGAC 783
56 |||||||
57
58 RESULT 30
59 US 09 908 1548 49
60 Sequence 39: Application US/099081548
61 Patent No. US20020168714A1
62 GENERAL INFORMATION:
63 APPLICANT: Barbas, Carlos F.
64 APPLICANT: Beegle, Roger
65 APPLICANT: Schopfer, Ulrich
66 TITLE OF INVENTION: REGULATION OF GENE EXPRESSION USING
67 TITLE OF INVENTION: SINGLE-CHAIN, MONOMERIC, LIGAND DEPENDENT POLYPEPTIDE
68 TITLE OF INVENTION: SWITCHES
69 FILE REFERENCE: TSKI 725.1
```



```

1  RESULT: 3
2  US: 06-9008 15408 4472
3  Sequence 44, Application: US/0900815408
4  Patent No.: US20020160871A1
5  GENERAL INFORMATION:
6  APPLICANT: Biorad, Catons P.
7  APPLICANT: Biorad, Catons P.
8  APPLICANT: Schopfer, Ulrich
9  TITLE OF INVENTION: REGULATION OF GENE EXPRESSION
10 TITLE OF INVENTION: SINGLE CHAIN, MOVEMENT
11 FILE REFERENCE: TSI 725, 1
12 CURRENT APPLICATION NUMBER: US/09/9008, 15408
13 PRIORITY FILING DATE: 2001-07-18
14 PRIORITY APPLICATION NUMBER: US 6/97619, 065
15 PRIORITY FILING DATE: 2000-07-18
16 NUMBER OF SEQ ID NOS: 54
17 SOFTWARE: FastSeq for Windows Version 4.0

```

```

1 RESULT 44
2 US 09/962,436-268
3 Sequence 268 Application US/09962436
4 Patent No. 55269200810101
5 GENERAL INFORMATION:
6 APPLICANT: SOPPOT, Daniel
7 TITLE OF INVENTION: Cancer Gene Determination and Therapies for Screening Using Stem
8 FILE REFERENCE: 689296-75
9 CURRENT APPLICATION NUMBER: US/09/962,436
10 PRIOR FILING DATE: 2001-09-25
11 PRIOR APPLICATION NUMBER: US/60/245,082
12 PRIOR FILING DATE: 2000-09-25
13 PRIOR APPLICATION NUMBER: US/60/234,924
14 NUMBER OF SEQ. ID NOS: 568
15 SOFTWARE: Patent in version 4.0
16 SEQ. ID NO 268
17 LENGTH: 2218
18 TYPE: DNA

```